

The use of mitochondrial DNA for the identification of fish in the early stages of development
from the eastern shores of Lake Huron, Ontario, Canada

by

Natalie D. J. Taylor

A thesis submitted in partial fulfilment
of the requirements for the degree of
Master of Science (MSc) in Biology

The Faculty of Graduate Studies
Laurentian University
Sudbury, Ontario, Canada

© Natalie D. J. Taylor 2016

THESIS DEFENCE COMMITTEE/COMITÉ DE SOUTENANCE DE THÈSE
Laurentian Université/Université Laurentienne
Faculty of Graduate Studies/Faculté des études supérieures

Title of Thesis Titre de la thèse	The use of mitochondrial DNA for the identification of fish in the early stages of development from the eastern shores of Lake Huron, Ontario, Canada	
Name of Candidate Nom du candidat	Taylor, Natalie	
Degree Diplôme	Master of Science	
Department/Program Département/Programme	Biology	Date of Defence Date de la soutenance November 07, 2016

APPROVED/APPROUVÉ

Thesis Examiners/Examineurs de thèse:

Dr. Douglas Boreham
(Supervisor/Directeur(trice) de thèse)

Dr. David MacLean
(Committee member/Membre du comité)

Dr. Frank Mallory
(Committee member/Membre du comité)

Dr. Mike Rennie
(External Examiner/Examineur externe)

Approved for the Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. Shelley Watson
Madame Shelley Watson
Acting Dean, Faculty of Graduate Studies
Doyenne intérimaire, Faculté des études
supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

I, **Natalie Taylor**, hereby grant to Laurentian University and/or its agents the non-exclusive license to archive and make accessible my thesis, dissertation, or project report in whole or in part in all forms of media, now or for the duration of my copyright ownership. I retain all other ownership rights to the copyright of the thesis, dissertation or project report. I also reserve the right to use in future works (such as articles or books) all or part of this thesis, dissertation, or project report. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that this copy is being made available in this form by the authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

Thesis Abstract

Ichthyoplankton specimens are notoriously difficult to identify using morphology. DNA barcoding and real-time PCR utilize DNA to identify specimens, rather than morphology. However, no study has yet compared morphological identifications with DNA-based identification for Canadian freshwater fishes. Here, we both compare DNA barcoding with morphological identification of ichthyoplankton and design a multiplex TaqMan real-time PCR assay for the rapid and cost-effective identification of 3 important species in Lake Huron, lake whitefish (*Coregonus clupeaformis*), deepwater sculpin (*Myoxocephalus thompsonii*) and round whitefish (*Prosopium cylindraceum*). Unlike morphological identification, DNA barcoding was able to resolve specimens from the *Catostomus* genus and *Cyprinidae* family to the species. Contrarily, DNA barcoding was unable to differentiate some members of the *Coregonus* genus, whereas morphology identified these specimens to the species. Real-time PCR was able to accurately identify the target species 100% of the time and was the most cost-effective method.

Keywords

DNA barcoding, TaqMan real-time PCR, morphological identification, lake whitefish, deepwater sculpin, round whitefish, larval fish

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr. Douglas Boreham for his guidance and support through this process. He has taught me valuable lessons that I will carry with me throughout my career. I would also like to thank Dr. Christopher Somers for his assistance in teaching me, and his patience in answering my many questions. Sorry for the nightmares that my project initially gave you. I would not have been able to complete my thesis without the help of these two mentors.

I would also like to express my gratitude to my lab-mates, in particular Andrew Zarnke who helped to teach me the ins and outs of real-time PCR and who was always available to help with troubleshooting. Thank you for listening to my ideas and providing me with a different perspective.

I would also like to thank my family for their endless support and encouragement. I would like to extend a special thank you to my brother, Jonathan, without whom I would never have considered pursuing a masters. Thank you also to my parents for taking an interest in my project (despite having no interest in fish or DNA) and encouraging me throughout my schooling.

Table of Contents

Thesis Abstract	iii
Acknowledgements	iv
List of Tables	vii
List of Figures	viii
List of Abbreviations	ix
Chapter 1: Introduction	1
Chapter 2: Hypotheses and Objectives	7
Chapter 3: DNA barcoding vs. morphological identification of larval fish and fish eggs in Lake Huron: advantages to a molecular approach	9
Abstract	10
Methods	14
Sample Collection	14
Morphological Identification	15
Molecular Identification	15
Data Analysis	17
Results	18
Morphological Identification	18
Molecular Identification	18
Morphological Identification vs. DNA Barcoding.....	20
Discussion	23
Deepwater Sculpin and Lake Whitefish	23
Disadvantages to the Morphological Approach	24
Disadvantages to DNA Barcoding	25
Identification of Fish Eggs	25
Advantages to DNA Barcoding	26
Conclusions	29
Chapter 4: The development of a novel multiplex TaqMan real-time PCR assay for the detection of lake whitefish (<i>Coregonus clupeaformis</i>), deepwater sculpin (<i>Myoxocephalus thompsonii</i>) and round whitefish (<i>Prosopium cylindraceum</i>)	30
Abstract	31
Introduction	32

Materials and Methods	36
Larval Fish Sample Collection.....	36
Reference Collection	36
DNA Extraction	37
DNA Barcoding	37
Sequence Selection.....	38
Primer and Probe Design and Verification/Validation	39
Results.....	43
Specificity	43
Application of Assay for Larval Fish and Fish Egg Identification	46
Cost of Real-Time PCR	46
Discussion	49
Acknowledgements	54
Chapter 5: Conclusions and Future Directions	55
References	60

List of Tables

Table 1 - The species and number of larval fishes that were morphologically identified and identified using DNA barcoding.

Table 2 - Comparing the morphologic and molecular identifications of larval fish, molecular identifications were considered to be correct. False positives indicate when a specimen was morphologically identified as the listed species, but was found to be another species using DNA barcoding. False negatives are specimens that were not morphologically identified as the indicated species, but were molecularly found to be the listed species.

Table 3 - The breakdown of cost for the molecular identification of larval fishes.

Table 4 - The primer and probe sequences for detection of species-specific DNA. Bases underlined with are bases that are variable between the target species and 1 other congeneric species, base highlighted with **G** are bases that are variable between the target species and 2 or more other congeneric species.

Table 5 - The average C_t of target DNA and congeneric DNA using the designed primers and probes. Maximum and minimum C_t values are the maximum and minimum over all runs and did not necessarily occur in the same test run. SE denotes standard error and N denotes the number of samples.

Table 6 - Cost breakdown for each component of the multiplex TaqMan real-time PCR. Cost per specimen was calculated based on the number of specimens that could be analyzed with each component.

List of Figures

Figure 1 - Diagram of a CANDU reactor depicting approximately where the larval fish and fish eggs sampling took place. Note: not to scale.

Figure 2 - Location of where the sampling took place on Lake Huron, Ontario, Canada.

Figure 3 – Box and whisker plot showing the mean, quartiles and maximum and minimum C_t values for all target and non-target species. Target species are in white, closely related species are in light grey, and other species are in dark grey.

Figure 4 - Example of a blind multiplex run where identifications were assigned by a blind technician and then confirmed with the identifications from DNA barcoding. Individual data points represent individual specimens.

List of Abbreviations

BLAST	Basic Local Alignment Search Tool
BOLD	Barcode of Life Database
bp	Base pairs
CAD	Canadian Dollar
COI	<i>Cytochrome Oxidase I gene</i>
COSEWIC	Committee on the Status of Endangered Wildlife in Canada
C _t	Cycle threshold
<i>Cyt b</i>	<i>Cytochrome b gene</i>
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphate
FAM	6-carboxyfluorescein
HEX	Hexachlorofluorescein
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
PCR	Polymerase Chain Reaction
T _m	Melting Temperature
Taq	<i>Thermus aquaticus</i>
VEC	Valuable Ecological Component

1 Chapter 1: Introduction

It is very important to accurately identify both adult and larval fish for a broad range of reasons, including: catching/releasing the appropriate species of fish in the appropriate season, food security, population management, and environmental controls. There are clearly defined criteria for the morphological identification of adult fish, which helps both amateur anglers and professional taxonomists accurately identify their fish. However, this is not always the case with ichthyoplankton. The study of ichthyoplankton is a growing field and the correct identification of larval fish and fish eggs is important for many different areas of research, such as providing information on the life cycles of poorly known species of fish, providing a better understanding of fish spawning habitat and migration, and assisting with the establishment of marine protected zones (1-5). There is also a demand from various industries to conduct more environmental studies to identify larval fish and fish eggs to help determine their potential environmental impact (6, 7). The correct identification of all species of fish is important for many different applications from conservation to industry, but there needs to be a confident method for the identification of ichthyoplankton.

The identification of fish in the early stages of development (larval and egg stages) are notoriously the two most difficult stages of development to accurately identify using morphological features (8, 9). Morphological identification relies on specific features that may not have yet developed in larval fish and fish eggs. Furthermore, in larval fish, when these features are present, they can be features that are similar amongst many species of fish at that stage of development, further compounding the difficulty of their proper identification (8, 10). The morphological identification of fish eggs often depends on the use of light microscopy to view the eggs (11). However, there are few distinguishing features on the surface of a fish egg

and as a result, fish eggs are often only able to be identified to the order or family (11). With this high potential for error and inaccuracy there has been a demand for an alternative method for the identification of larval fish and fish eggs (11-14).

DNA barcoding has been used as an alternative method to morphological identification of larval fish and fish eggs (10, 15, 16). This method relies on the PCR amplification and subsequent sequencing of *Cytochrome oxidase subunit I* (COI), a 658 base pair gene in the mitochondrial genome in animal species (17, 18). The obtained genetic sequence is then compared to a database of reference COI sequences. Originally used for the identification of microbes, DNA barcoding became popular in 2003 when the COI gene was suggested as the standard target barcode gene for all animal species (19). In principle, DNA barcoding relies on the “barcode gap” - the difference between interspecific variability and intraspecific variability (18-20). In order to be able to differentiate two species based on their COI sequences, there has to be a high level of interspecific variability, to allow for differentiation between species; there also has to be a low level of intraspecific variability, so that individuals from the same species do not differ on a genetic level. The mitochondrial genome was chosen as the target, rather than the nuclear genome, because it is highly conserved within species but still provides adequate variability between different species (21-23). Furthermore, mitochondrial DNA is present in the cell in multiple copies, there are no introns, and there are few duplications, making it an ideal target for DNA barcoding (23, 24). DNA can identify individual specimens to the species level, and thus provides an alternative method to morphological identification with the potential for higher accuracy.

Although the COI gene has been widely used, and used successfully to identify fish, there are also disadvantages associated with COI, due to the nature of the mitochondrial genome. One

disadvantage of using mitochondrial DNA for species identification is that it is transmitted through maternal inheritance and therefore does not allow for the detection of hybrid species (e.g. a fish with a walleye [*Sander vitreus*] mother and sauger [*Sander canadensis*] father would be identified as walleye using DNA barcoding) (23, 25). Secondly, with DNA barcoding there are some genera in which there is inadequate interspecific differentiation for the differentiation of some of these species, meaning that these individuals would only be identified to the genus level. April et al. (2011) identified some North American species of fish that cannot be differentiated using the COI gene, as they do not have enough interspecific variability (26). However, DNA barcoding can still identify approximately 90% of North American freshwater species of fish (26). Finally, a comprehensive database of COI sequences is required to create a reference group with which to compare new sequences (21, 26, 27). If no such database exists, it is possible that the target species of fish is not present in the database, which can result in erroneous identifications. However, for Canadian freshwater fishes, there are comprehensive databases that have been established (28, 29). DNA barcoding provides a quick and reliable way to identify most fishes but as a result of the maternal inheritance, and potential weak levels of interspecific differentiation and lack of a comprehensive database for some geographic regions, the COI gene may not be the only appropriate gene for use as a universal barcode region.

While both DNA barcoding and morphology have been used to identify larval fish and fish eggs, DNA barcoding has been shown to be 99-100% accurate (depending on the size of the database being used for comparison) (27). However, there have been few studies that compare which of these two methods is more accurate and cost-efficient and there have been no studies comparing the two methods in Canadian freshwater fish.

Real-time PCR has also been used to identify fish species. There are 2 general forms of real-time PCR, which are classified based on the detection method. Intercalating fluorescent dyes, like SYBR Green, can be used for the detection and quantification of double stranded DNA (30, 31). While it is possible to run multiple samples simultaneously through multiplexing with SYBR Green, this requires gel electrophoresis to compare product size (as SYBR Green cannot differentiate products during the real-time PCR amplification) (32, 33). TaqMan real-time PCR is more appropriate for multiplexing applications as it utilizes multiple different fluorescent-labelled probes that are designed to bind the target sequence between the forward and reverse primers. The different fluorophores can be individually detected, while in the same reaction, thus eliminating the need to run a gel to differentiate products based on size. TaqMan real-time PCR relies on the 5'-3' exonuclease activity of *Taq* Polymerase (34). When the probe is intact, the unique wavelength of fluorescence emitted by the fluorophore (on the 5' end of the probe) is masked by a quenching molecule (on the 3' end of the probe) (35). However, when *Taq* polymerase amplifies the target DNA, it digests the probe, thereby releasing the fluorophore and quencher from the probe. As the masking action of the quencher is dependent on being in close proximity to the fluorophore, when the probe is digested they separate and the fluorescence is detected and can be quantified for objective analysis (35).

TaqMan real-time PCR is ideal for multiplexing applications because different probes can be customized with different fluorophores. Different fluorophores have different excitation and emission spectra that can be differentiated by the real-time PCR machine allowing for the differentiation of various target sequences from one another in the same reaction. TaqMan real-time PCR has been used for a wide breadth of applications including the detection of various pathogens in medicine and agriculture, the verification of meat product identities in the food

service industry and the differentiation of microbes (36-40). Although the TaqMan approach has been used in a wide variety of organisms with high levels of multiplexing, it has not been used in its multiplex form for the identification of larval fish and fish eggs of Canadian freshwater fishes.

Here, TaqMan PCR was used to differentiate 3 important species of fish, namely lake whitefish (*Coregonus clupeaformis*), deepwater sculpin (*Myoxocephalus thompsonii*) and round whitefish (*Prosopium cylindraceum*). These three species of fish have been identified as valuable ecological components (VECs) by various nuclear power plants on the Laurentian Great Lakes, like Bruce Power – a nuclear power plant on the east shores of Lake Huron (41, 42). Lake whitefish has been identified as a VEC because of its value to the commercial and Native fishing industry on Lake Huron. These fish are harvested for both their meat and roe and in 2016 represented the third largest commercial fishery in Ontario by landings (43, 44). Lake whitefish are important both culturally and economically in Canada, making it important to carefully monitor and research this valuable species to understand the potential industrial impacts.

Deepwater sculpin live at the bottom of deep, cold, well-oxygenated lakes. They are a VEC in Lake Huron because they are classified as a species of special concern in Ontario by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (41, 45). Deepwater sculpin are an important prey for other species of fish such as burbot (*Lota lota*) and lake trout (*Salvelinus namaycush*) and are therefore an important part of the profundal food chain (45). As a result of their lower position on the food web, they are considered to be an indicator of the health of deepwater fish communities (45). Deepwater sculpin have already been extirpated from Lake Ontario and Lake Erie increasing the importance of continued monitoring by industries like Bruce Power (46). Research into the larval and egg stages of the deepwater

sculpin can provide valuable information for researchers into the life cycle of this poorly understood species of fish.

Round whitefish have previously been studied by industries, like Bruce Power, that use lake water from the Laurentian Great Lakes in their operations. It has been suggested that the thermal plume generated by nuclear power plants may have an impact on round whitefish, as they may be more thermally sensitive than other species of fish, such as lake whitefish (47, 48). They are an important species to monitor, as they are known to inhabit areas around several nuclear power plants on the shores of the Laurentian Great Lakes (41, 42).

This project will be useful for both research and industry based on the Laurentian Great Lakes and in the remainder of Canada. Through comparing morphological and molecular identification techniques and designing a novel multiplex assay to accurately identifying ichthyoplankton to the species level, we hope that this will provide both industries and researchers with the tools that are needed to better understand larval fish in terms of their abundance and dynamics. While this research was targeted for use in environmental studies, we believe that it will also be useful for other applications such as investigations into the life cycles of fish, food safety and fish population management.

2 Chapter 2: Hypotheses and Objectives

The morphological identification of fish can be a subjective procedure and the results from identification efforts of the same specimens by different taxonomists can result in different identifications, especially in larval fish (10). Morphological identification is dependent on the gross anatomy of the specimen and is therefore dependent of the entire specimen remaining physically intact. DNA barcoding is an objective procedure that depends primarily on the quality of DNA, not the gross integrity of the specimen, and therefore, is still a viable option if the specimen has sustained some sort of physical damage. Similarly, TaqMan real-time PCR also relies on the quality of the DNA rather than the gross physical characteristics of fish. It is a rapid identification tool that can be customized for the detection of any target species. Therefore, the aim of this thesis is to compare morphological identification with molecular identification, namely, DNA barcoding for species identification of larval fish, and to design an identification tool for the rapid identification of VECs.

Hypothesis: DNA barcoding will be able to provide greater species resolution compared to morphological identification techniques. A novel multiplex TaqMan real-time PCR assay will be able to detect and differentiate DNA from species of fish that are VECs. This assay will be a more cost-effective option compared to both DNA barcoding and morphological identification.

The hypothesis will be evaluated by through the following experimental objectives:

Objectives:

1. Identify larval fish using morphology.
2. Identify larval fish and fish eggs using DNA barcoding.
3. Identify candidate sites for primers and probes on the COI and *Cyt b* genes for real-time PCR.
4. Validate the use of designed primers and probes in a multiplex TaqMan real-time PCR assay.
5. Compare the cost of each method on a per-fish basis

3 Chapter 3: DNA barcoding vs. morphological identification of larval fish and fish eggs in Lake Huron: advantages to a molecular approach

**DNA barcoding vs. morphological identification of larval fish and fish eggs in Lake Huron:
advantages to a molecular approach**

(Original Research)

Natalie D.J. Taylor, Christopher M. Somers, Richard G. Manzon,
Joanna Y. Wilson, Douglas R. Boreham

[In review for publication in the Journal of Great Lakes Research]

3.1 Abstract

Reliable identification of larval fish and eggs is essential for many types of ecological studies, including environmental monitoring. DNA barcoding may be an effective alternative to morphological identification, but few studies have compared the two approaches. We identified 657 larval fish representing 9 families from Lake Huron (Ontario, Canada) both morphologically and using DNA barcoding at the cytochrome oxidase I locus (COI). We used DNA barcoding only to attempt identification of 103 fish eggs. For larval fish that were successfully identified using both methods, agreement between the two methods was 76.9%, 96.6% and 96.6% at the species, genus, and family levels respectively. Damaged specimens resulted in 37 (5.6%) failed identifications (unknowns) using morphology; 35 of these specimens were successfully identified using DNA barcoding. However, 23 (3.8%) other specimens produced no PCR product for barcoding using 2 different primer sets. We were able to identify 52 (50.5%) of the fish eggs using DNA barcoding. Discrepancies between morphology and DNA barcoding for larval fish were driven largely by 3 major factors: (1) inability of COI to resolve members of the genus *Coregonus*; (2) limited resolution of morphological features for *Catostomus* and Cyprinidae; and (3) a variety of mismatches affecting 21 (~4%) other mismatched specimens. Our findings ultimately demonstrate equivocal overall performance of the two techniques; however, DNA barcoding has the added advantages of being faster, cheaper, and requires less specialized training.

Keywords: DNA barcoding, morphological identification, ichthyoplankton.

3.2 Introduction

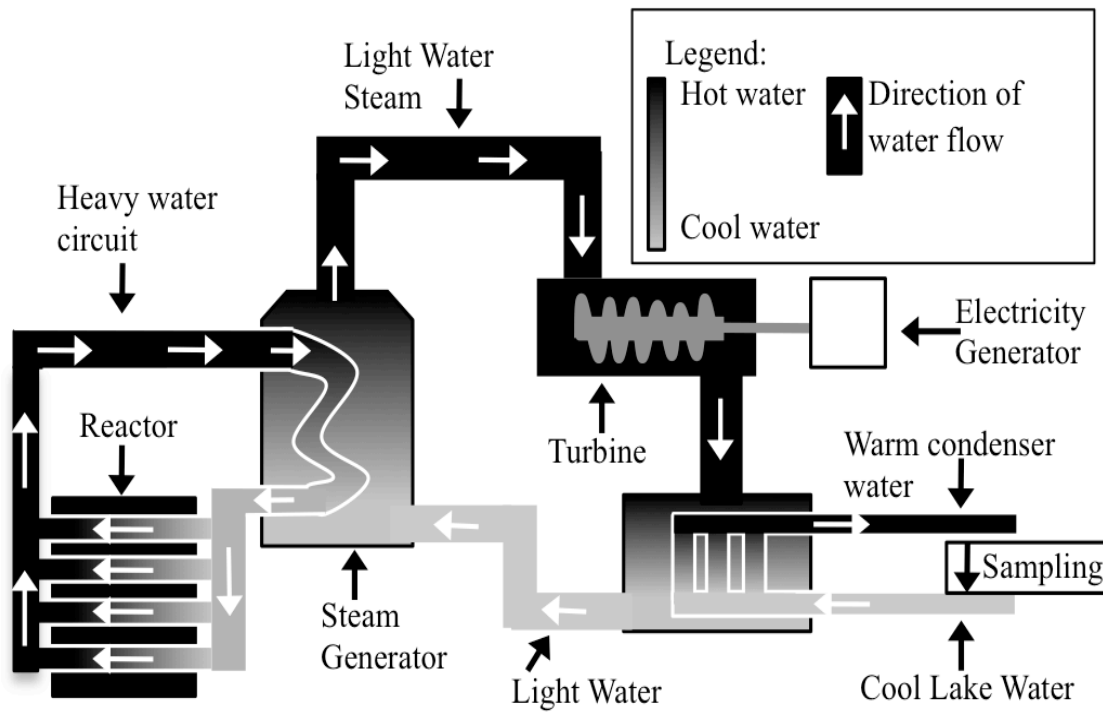
Correct identification of larval fish and fish eggs is important for a broad range of research, including the identification of fish nursery grounds and dispersal routes, the prevention of illegal trade, and the establishment of protected marine zones (10, 49-51). The identification of larval fishes based on morphology can be difficult for several reasons. Many species share very similar traits at the larval stage, which makes accurate resolution difficult (10). In contrast, the larval stage in some species may appear quite different among members of the same species based on relative age (8, 9). Consequently, the early life stages of fish are the most difficult times to identify to species level. This problem is more challenging when larval fish specimens are damaged by industrial processes, or for fish in the egg stage, which generally cannot be identified morphologically (8, 9). As a result, any attempt at morphological identification requires highly trained taxonomists, and even with this training, it is recommended that larval fish only be identified to the family level (10). However, ecological studies often require information at the genus or species level. Thus, there is a need for an alternative method to supplement or replace the traditional morphological approach.

DNA barcoding has been widely used for the identification of many diverse species (13, 16, 19, 52, 53). DNA barcoding uses the cytochrome *c* oxidase subunit I (COI) region of the mitochondrial genome for species identification (19, 26, 54-56). The basis for species differentiation using DNA barcoding is that there is a higher degree of interspecific sequence variation than intraspecific variation in this region (18, 27, 54). An advantage of DNA barcoding is that it can be used to identify the species of virtually any fish sample, no matter the life-stage or level of physical damage, as long as DNA can be recovered from the sample for PCR (54, 57, 58). As such, DNA barcoding has been widely used for the identification of larval fish (16, 57,

58); however, there have been few studies comparing morphological identification done by expert taxonomists with DNA barcoding to determine which is the more appropriate technique (10, 15).

Here we compare the morphological identification of larval fish collected from the water intake system at Bruce Power, a large CANDU nuclear facility, on Lake Huron, Ontario, Canada, with DNA barcoding of the same specimens. We used DNA barcoding to assess the identity of individual fish eggs. It is important to understand which larval fish and eggs are being drawn into the water intake system for the cooling of steam condensers at Bruce Power (Figure 1) as part of ongoing research into potential environmental and fisheries impacts associated with energy generating stations that use natural lake, river and sea water as industrial coolant (6, 59). Overall, our objective was to establish which method of identification was better for long-term environmental studies, but we were also specifically interested in the accuracy of morphological identifications of lake whitefish (*Coregonus clupeaformis*) and deepwater sculpin (*Myoxocephalus thompsonii*). Lake whitefish is ecologically and culturally important in Lake Huron, supporting a large commercial fishery (41, 60, 61). Deepwater sculpin is listed as a species of special concern because of declining populations in the Laurentian Great Lakes (62).

Figure 1 - Diagram of a CANDU reactor depicting approximately where the larval fish and fish eggs sampling took place. Note: not to scale

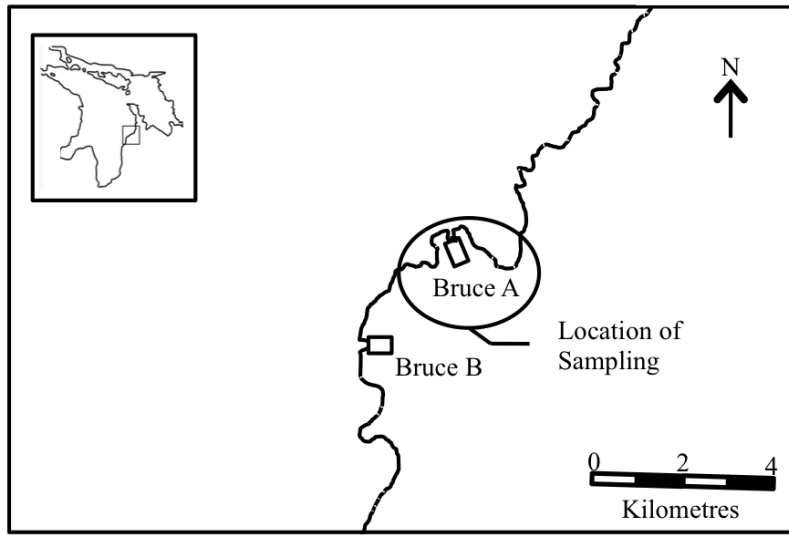


3.3 Methods

3.3.1 Sample Collection

Larval fish and fish egg specimens were collected from March through December in 2013 and 2014 from the Bruce Power water intake on the east shore of Lake Huron near Tiverton, Ontario, Canada (N 44.338°, W 81.573°) (Figure 2). Sampling took place at the Bruce A facility using a round, 500µm mesh plankton net that was 0.72m in diameter and was lowered 3m into the intake water for 5 to 140 minutes (median 36 minutes). Specimen collection occurred both during the day and at night. Between 2013 and 2014, there were 81 day and 80 night sampling efforts, with a minimum of 3 nets set per effort. Each specimen collected was given a unique identification number and larval fish were stored in 95% ethanol until morphological analysis. After morphological analysis, larval fish of the same species from the same collection time were stored together in 95% ethanol. Fish eggs from the same collections were stored together in 95% ethanol. All efforts resulted in the collection of 1740 larval fishes and 2831 fish eggs. Larval specimens were chosen for both morphological identification and DNA barcoding so as to capture a representative sample from all possible species. Fish eggs were randomly selected for analysis.

Figure 2 - Location of where the sampling took place on Lake Huron, Ontario, Canada



3.3.2 Morphological Identification

Samples were analysed commercially (through an external contractor) by an expert ichthyologist who specializes in the identification of larval fishes from the Laurentian Great Lakes. Specimens were identified based on body shape, myomeres, pigmentation, meristic count, and fin characteristics etc. (e.g., number, shape, relative position etc.). When possible, specimens were identified to the species level; otherwise, specimens were identified to the genus or family level. Results of morphologic identifications were recorded based on the unique identification number of each specimen. Fish eggs were not identified morphologically. The cost of these identifications was recorded for a comparison with DNA barcoding. Cost was calculated on a per-specimen basis in US Dollars (USD) and converted to Canadian Dollars (CAD).

3.3.3 Molecular Identification

657 larval fish and 103 fish eggs were selected for molecular identification from all samples collected. The 657 larval fish were selected to ensure a representative sample, and to ensure that

there were members from all potential species of fish. 657 was deemed to be an appropriate amount for our analysis. Only 103 fish eggs had DNA extracted due to the difficulty associated with DNA extraction from fish eggs and the high rate of failure.

DNA was extracted from individual larval fish and eggs using spin column kits according to manufacturer guidelines (Qiagen DNEasy, Mississauga, ON; Norgen Biotech DNA extraction kit, Thorold, ON). When larval specimens were small (<12mm in total length), the entire fish was used for DNA extraction; when larval specimens were larger, a portion of the body (up to 12mm) was used for DNA extraction. DNA concentration from extractions was measured using a NanoDrop 1000 spectrophotometer or Qubit fluorometric quantitation (Life Technologies).

For all specimens, a 658 bp region of the COI mitochondrial genome was PCR-amplified using universal primers FishF1 (5'-TCA ACC AAC CAC AAA GAC ATT GCC AC-3') and FishR1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3') (18). PCR reactions consisted of a total volume of 25µL with components as follows: 1x PCR buffer; 2.5mM of MgCl₂; 0.1µM of each forward and reverse primers, 0.05mM of each dNTPs, 0.3125 units of Taq DNA polymerase, and 10ng of template DNA. The thermal cycling regime consisted of: 2 minutes at 94°C followed by 35 cycles of: 30 seconds at 94°C, 40 seconds at 52°C and 1 minute at 72°C. Final extension was for 10 minutes at 72°C. PCR products were run on a 1% Agarose gel to verify the presence of a product in the target size range. Specimens that failed the initial PCR were run a second time using the universal fish primers FF2d (5'-TTC TCC ACC AAC CAC AAR GAY ATY GG-3') and FR1d (5'-CAC CTC AGG GTG TCC GAA RAA YCA RAA-3') (56) using the reaction conditions specified above. Negative control samples (with no template) were randomly dispersed throughout the PCR runs to ensure that there was no contamination.

Successfully amplified DNA was purified using the Qiagen MinElute PCR Purification Kit (Mississauga, Ontario) with a final elution volume of 10 μ L. Final elution was performed using MilliQ biology-grade water or nuclease-free water; 3.2pmol of forward primer (FISH F1 or FF2d) was added to the purified product and samples were Sanger sequenced (University of Calgary, Core DNA Services). The overall cost for the identification of the specimens using DNA barcoding was recorded for later comparison. Cost was calculated on a per specimen basis, in CAD and included labour (for a graduate student), price for laboratory supplies used and sequencing costs.

3.3.4 Data Analysis

Sequences were aligned and compared using SeaView V. 4.5.4 (63). Specimens were identified by comparing COI sequences obtained against those in the NCBI database using the Basic Local Alignment Search Tool (BLAST) (64). We used quantitative criteria similar to those described by Ko et al. (2013) for identifications. Specifically, specimens were identified to the species level when they matched an individual species at >98% similarity and the maximum bit score returned by BLAST was over 1000, with the bit score of the next most likely species match lower by more than 100 bit points. When a clear top match was not present, i.e., when there was more than one species with >98% match, or no species >98% match, samples were identified to the top-matching genus or family. Specimens were identified to the genus and family when the similarity values were 97%-92% and 91%-84%, respectively. Geographic distribution of fish species was taken into account when identifying the fish specimens; i.e., fish species that do not occur in the study region were not included for consideration.

3.4 Results

3.4.1 Morphological Identification

Of the 657 larval fish that were analysed, 620 were identified to family (94.4%), 618 to genus (94.1%), and 583 to species (88.7%). Of the total specimens analysed, 37 (5.6%) could not be identified because they were damaged in some way (i.e., missing key features for identification). According to morphological identification, only burbot (*Lota lota*), deepwater sculpin (*Myoxocephalus thompsonii*), bloater (*Coregonus hoyi*), rainbow smelt (*Osmerus mordax*), and round goby (*Neogobius melanostomus*) each made up more than 5% of the total specimens analyzed (Table 1). Burbot and deepwater sculpin were at least two times more abundant than the next closest species (bloater).

3.4.2 Molecular Identification of Larval Fish

Of the 657 larval fish specimens analyzed using COI sequence data, 632 were identified to family (96.1%), 632 to genus (96.1%), and 532 to species (81.0%). 25 specimens (3.8%) could not be identified due to amplification failure with both sets of PCR primers, 23 of which were identified using morphology. However, of the 37 specimens that could not be identified morphologically, 35 were identified to the species level using DNA barcoding.

Table 1 - The species and number of larval fishes that were morphologically identified and identified using DNA barcoding

Family	Genus	Species	#ID Morphology (%)	#ID COI (%)
<i>Clupeidae</i>	<i>Alosa</i>	<i>pseudoharengus</i> ¹	1 (0.2)	1 (0.2)
<i>Catostomidae</i>	<i>Carpoides</i>	<i>cyprinus</i> ²	2 (0.3)	3 (0.5)
	<i>Catostomus</i>	Genus only	20 (3.0)	0
		<i>commersonii</i> ³	0	17 (2.6)
		<i>catostomus</i> ⁴	0	1 (0.2)
<i>Salmonidae</i>	<i>Coregonus</i>	Genus only	15 (2.3)	100 (15.2)
		<i>artedi</i> ⁵	27 (4.0)	0
		<i>hoyi</i> ⁶	60 (9.1)	0
		<i>clupeaformis</i> ⁷	7 (1.0)	4 (0.6)
	<i>Oncorhynchus</i>	<i>tshawytscha</i> ⁸	1 (0.2)	0
		<i>mykiss</i> ⁹	0	1 (0.2)
<i>Cyprinidae</i>	Family only	Family only	2 (0.3)	0
	<i>Notropis</i>	<i>atherinoides</i> ¹⁰	0	5 (0.8)
		<i>hudsonius</i> ¹¹	0	1 (0.2)
	<i>Cyprinus</i>	<i>carpio</i> ¹²	0	1 (0.2)
<i>Lotidae</i>	<i>Lota</i>	<i>lota</i> ¹³	243 (37.0)	257 (39.1)
<i>Cottidae</i>	<i>Myoxocephalus</i>	<i>thompsonii</i> ¹⁴	144 (21.9)	147 (22.4)
<i>Gobiidae</i>	<i>Neogobius</i>	<i>melanostomus</i> ¹⁵	43 (6.5)	44 (6.7)
<i>Osmeridae</i>	<i>Osmerus</i>	<i>mordax</i> ¹⁶	48 (7.3)	43 (6.5)
<i>Percidae</i>	<i>Perca</i>	<i>flavescens</i> ¹⁷	6 (0.9)	6 (0.9)
	<i>Sander</i>	<i>vitreus</i> ¹⁸	1 (0.2)	1 (0.2)
		Unidentified	37 (5.6)	25 (3.8)
		Total	657	657

Common Fish Names: 1 – Alewife, 2 – Quillback, 3 – White Sucker, 4 – Longnose Sucker, 5 – Cisco, 6 – Bloater, 7 – Lake Whitefish, 8 – Chinook Salmon, 9 – Rainbow Trout, 10 – Emerald Shiner, 11 – Spottail Shiner, 12 – Common Carp, 13 – Burbot, 14 – Deepwater Sculpin, 15 – Round Goby, 16 – Rainbow Smelt, 17 – Yellow Perch, 18 – Walleye

DNA barcoding was able to identify lake whitefish (*Coregonus clupeaformis*), but could not differentiate between cisco (*Coregonus artedi*), kiyi, (*Coregonus kiyi*) and bloater. The 100 specimens (17% of sample population) that were either cisco, kiyi or bloater were therefore only identified to the genus. According to DNA barcoding, species exceeding 5% of our collections

were: burbot, deepwater sculpin, the genus *Coregonus* (excluding lake whitefish), rainbow smelt, and round goby (Table 1).

3.4.3 *Molecular Identification of Fish Eggs*

We were able to amplify a PCR product from 52 (50.5%) of the individual fish eggs collected, of the 103 eggs from which DNA was extracted. All 52 (100%) of the eggs identified were walleye (*Sander vitreus*). These eggs were from 14 different collection efforts that took place in April (2 collections), May (5 collections) and June (2 collections) in 2013, and in May (4 collections) and June (1 collection) in 2014. COI data for walleye eggs contained several different haplotypes across the sampling periods, suggesting that egg intake is common and includes a variety of females.

3.4.4 *Morphological Identification vs. DNA Barcoding*

There were 3 major points of difference between identifications made using the two approaches. The first, and most important in terms of numbers, was that DNA barcoding could not differentiate three members of the genus *Coregonus* (specifically, bloater, kiyi and cisco), which resulted in 100 (16.8%) discordant identifications to the species, of the 597 specimens identified using both techniques. Second, DNA barcoding was capable of resolving suckers (*Catostomus*) and Cyprinids to species, whereas morphology was limited to the genus or family level, leading to 24 (4.0%) discordant identifications at the species level of the identified specimens (Table 1). Finally, there were a remainder of 20 (3.4%), 20 (3.4%), and 21 (3.5%) sporadic discordant identifications to the family, genus and species respectively, excluding the specimens that were unidentified (Table 2). The most consistent disagreement when both techniques identified specimens to the species was caused by the lake whitefish, for which 5 of 7 identifications did

not match. In 4 of the 5 cases, DNA barcoding revealed that other *Coregonus* species (either bloater, kiyi or cisco) were incorrectly identified as larval lake whitefish (false positives). The remaining misidentification was a lake whitefish that was morphologically identified as a member of the *Catostomus* genus (false negative). There was also disagreement for rainbow smelt; 5 of 48 specimens morphologically identified as rainbow smelt were actually burbot, emerald shiner (*notropis antherinoides*), or yellow perch (*perca flavescens*) according to barcoding. Two specimens were morphologically identified as burbot, but were identified as rainbow smelt using DNA barcoding. In contrast, only 3 of 144 deepwater sculpin identifications disagreed; 1 specimen was morphologically identified as a bloater, and 2 as burbot.

Table 2 - Comparing the morphologic and molecular identifications of larval fish, molecular identifications were considered to be correct. False positives indicate when a specimen was morphologically identified as the listed species, but was found to be another species using DNA barcoding. False negatives are specimens that were not morphologically identified as the indicated species, but were molecularly found to be the listed species.

Species/Genus/Family	False positives	False Negatives
Quillback (<i>Carpoides cyprinus</i>)	0	1
Sucker (<i>Catostomus spp.</i>)	2	0
Whitefish (<i>Coregonus spp.</i>)	2	6
Bloater (<i>Coregonus hoyi</i>)	4	0
Lake Whitefish (<i>Coregonus clupeaformis</i>)	4	1
Carp/Minnow Family (<i>Cyprinidae</i>)	2	0
Common Carp (<i>Cyprinus carpio</i>)	0	1
Burbot (<i>Lota lota</i>)	7	6
Deepwater Sculpin (<i>Myoxocephalus thompsonii</i>)	0	3
Round Goby (<i>Neogobius melanotsomus</i>)	0	1
Chinook Salmon (<i>Oncorhynchus tshawytscha</i>)	1	0
Rainbow Trout (<i>Oncorhynchus mykiss</i>)	0	1
Rainbow Smelt (<i>Osmerus mordax</i>)	5	2
Yellow Perch (<i>Perca flavescens</i>)	0	1
Emerald Shiner (<i>Notropis atherinoides</i>)	0	3
Spottail Shiner (<i>Notropis hudsonius</i>)	0	1
Total	27	

In terms of cost, the external contractor charged \$10 USD to identify each larval fish (~13.30 CAD) while DNA barcoding cost \$12.68 CAD per specimen (Table 3), leading to a difference in cost of \$0.62 per larval fish. Note: the labour used for this calculation was that of a graduate student, and the cost would increase should another form of human labour be used, however, it could be reduced over time if a robot were to be used instead.

Table 3 - The breakdown of cost for the molecular identification of larval fishes

Material Used	Total Cost	Cost per Specimen
Qiagen DNEasy Blood and Tissue Kit	\$1,511.10	\$2.30
Qiagen Taq PCR Core Kit	\$525.60	\$0.80
Primers	\$20.20	\$0.03
Agarose gel and Loading Dye	\$118.26	\$0.18
Qiagen MinElute PCR Purification Kit	\$1,149.38	\$1.75
Sequencing	\$3,942.00	\$6.00
Labour	\$1,064.34	\$1.62
Totals	\$8,330.88	\$12.68

3.5 Discussion

Morphological identification and DNA barcoding both performed reasonably well for identification of larval fish, and our data do not clearly demonstrate which is the better approach. The success of morphological identifications depends on the life-stage and quality of the specimen, while the success of DNA barcoding depends on the degree of interspecific variation (55, 60). DNA barcoding has been shown to be 99-100% accurate when there is a comprehensive database for comparison, which is mostly the case for Canadian freshwater fishes (27).

3.5.1 *Deepwater Sculpin and Lake Whitefish*

Deepwater sculpin (species of Special Concern) and lake whitefish (commercial fishery) identifications were of particular interest in our study. For deepwater sculpin there was only a 2% discrepancy rate between techniques. This demonstrates that DNA barcoding and morphological identification can both be used reliably for this species. A possible reason for this high level of accuracy is the lack of closely related species in Lake Huron, which would make it more difficult to differentiate among members of the same genus (65-67). In contrast, for lake whitefish, most of the discrepancies we identified were morphological false positives, which would lead to an overestimation of the number of lake whitefish larvae in the sample population using this method. False positives are important to avoid in environmental monitoring programs because they can lead to inappropriate measures to reduce ecological impacts that may not actually exist. The equivocal performance of the two techniques for deepwater sculpin, and the potential for inaccurate identification of lake whitefish using morphology, together suggest that DNA barcoding is the better method for studies examining these two species.

3.5.2 Disadvantages to the Morphological Approach

For the Cyprinidae family, morphology was unable to identify any of the specimens beyond the family, while DNA barcoding was able to identify these specimens to the species. Morphology also frequently misassigned members of this family to other families. DNA barcoding is able to differentiate the majority of species in this family from one another due to the high level of interspecific variability in the COI region within this family. April et al. (2011) found that, of the 221 species they analyzed from this family, only 10 species could not be differentiated using DNA barcoding. It is unknown as to why the morphological approach was not able to identify these specimens beyond the family as there are defined criteria for their delineation (68).

For the *Catostomus* genus, DNA barcoding is clearly able to identify the specimens further than using morphology, as morphology only identified these specimens to the genus and DNA barcoding resolved them to the species. DNA barcoding is able to easily differentiate longnose sucker (*Catostomus catostomus*) from white sucker (*Catostomus commersonii*) due to the variability in their COI sequences. Our analysis showed that there were 52 consistently variable base pairs between these two species, making their differentiation using DNA barcoding clear and simple. Contrarily, there was great difficulty to resolve these species beyond the genus using morphology. This is not a difficulty that has been limited to this study. It has been noted that morphology cannot differentiate longnose and white sucker until scales have formed on the specimens (in the juvenile stage of development) (69, 70). This difficulty can partially be attributed to their lack of difference in their pigmentation and the fact that their myomeres are too similar at all stages of early development (71).

3.5.3 Disadvantages to DNA Barcoding

A major problem for potential Great Lakes applications is that DNA barcoding with COI is unable to differentiate bloater (*Coregonus hoyi*), kiyi (*Coregonus kiyi*) and cisco (*Coregonus artedii*). COI sequences have no consistent variation between these three species, making their differentiation using COI barcoding impossible. Consequently, these specimens in our collection could only be identified to the genus level using the molecular approach. This is a persistent problem with the *Coregonus* genus (26, 54, 61). In 2011, April et al. showed that there were 7 species in the *Coregonus* genus that were indistinguishable when DNA barcoding was used (including bloater, kiyi and cisco). Bloater, kiyi and cisco specimens comprised 17% of the collection we assessed, so members of the *Coregonus* genus had a major influence on our perception of the performance of DNA barcoding. However, there is no means of confirming the accuracy of morphological identification of bloater and cisco, so we cannot truly evaluate the performance of the two techniques for this difficult group. This challenge with COI barcoding may be fairly widespread for fish. April et al. (2011) found 75 species of 752 species analyzed (10%) that could not be identified through DNA barcoding of the COI gene. It is therefore important that a different locus be established and verified for enhanced resolution, such as the 16S subunit, ITS1 or Cytochrome *b* regions of the genome that have been used previously for other application (e.g. Real-Time PCR) (72, 73). This could allow for multi-locus DNA barcoding, thus ensuring that all specimens may be accurately identified to the species.

3.5.4 Identification of Fish Eggs

We were able to generate DNA barcode data from individual fish eggs in over half of the analysed samples. All of the eggs were identified as walleye, which are a valuable commercial

and recreational species. Multiple spawning females would have deposited eggs in an area of the lake near the water intake system, or been physically impinged while gravid. This information may allow for additional considerations for walleye at this location. However, in assessing possible reparation measures, it is important to remember that the number of eggs analyzed (103) was a very small subset of the collected egg population (2381) and that further analysis needs to be done before drawing any conclusions. The identification of fish eggs in general will provide industries that utilize lake, river and seawater with a more accurate assessment of their environmental impact.

3.5.5 Advantages to DNA Barcoding

Another factor to take into consideration is the accuracy of both techniques in terms of identification at the species level. When DNA barcoding was used to validate morphological identification in this study, we found that there were morphological misidentifications to the species level 3.5% of the time (a total of 21, only specimens with species-level identifications for both techniques were included). While this demonstrates that morphological identifications (by an expert ichthyologist) are highly accurate, it shows that DNA barcoding is a slightly more accurate technique, when DNA barcoding is able to identify the specimen to the species (with an accuracy of 99-100%) (10, 27). Our study shows that lake whitefish, rainbow smelt, and burbot are the most likely to be misidentified using morphology (71%, 16%, and 5% misidentification rates respectively). The difficulty associated with morphological identification to the species can have serious consequences on study results. When there are many false positives, as with lake whitefish, rainbow smelt and burbot, this can lead to an overestimation of the ecological impact on these species, and an underestimation of the impact on other species. For environmental impact studies, these misidentifications can lead to erroneous repopulation efforts, and erroneous

controls and quotas on these species etc., all of which can result in the mismanagement of the fish population.

DNA barcoding was able to identify a slightly higher percentage of specimens (96.2%) than morphological identification (94.4%) when all taxonomic levels were considered. This is because DNA barcoding does not rely on the physical quality of the specimen appearance and can accurately identify larval fish independent of the life-stage or specimen condition (10). As such, barcoding is able to identify specimens that have been damaged, are missing key diagnostic features, or that are in stages of development that cannot be identified using morphology (10). All larval fish specimens that could not be identified morphologically were damaged in some way; e.g., some specimens were not complete, others were not preserved appropriately, etc. However, DNA barcoding was able to identify 95% of these specimens, demonstrating the versatility of DNA barcoding. As such, DNA barcoding should be used when the specimen is damaged in such a way as to limit morphological identification. However, when the specimen is not damaged, our study shows that both morphological identification and DNA barcoding are robust means of larval fish identification.

In terms of cost and time invested, the external contractor charged \$10 USD to identify each larval fish (~\$13.30 CAD), whereas DNA barcoding, as we ran it, cost \$12.68 CAD per specimen. The cost of DNA barcoding could be lowered considerably (by 50% or more) for high-throughput situations, and by seeking out lower sequencing costs (as sequencing in this study cost \$6.00 CAD per specimen). In addition, a graduate student could turn around DNA barcoding data as presented here in days to weeks without specialized training, as DNA barcoding required ~20 hours of training, whereas it takes years to become an expert larval or egg ichthyological taxonomist. Finally, damaged specimens can still provide useful DNA for

barcoding, even when they are missing all identifying morphological features. Thus, we conclude that although the performance of the two techniques is very similar in terms of the data produced, DNA barcoding offers some distinct advantages that make it a potentially more attractive option in some cases. The capacity for high-throughput DNA barcoding would allow for a faster, cheaper (when conducted by graduate students) and more efficient method of larval fish identification. However, for the identification of offshore coregonids, a combination of morphological identification and DNA barcoding should be used to ensure accurate identification (as each method struggles to accurately identify different species from this genus).

3.6 Conclusions

DNA barcoding and morphological identification provide equivocal performance overall in terms of resolution. However, DNA barcoding has the advantage of being cheaper (when graduate students are used as the labour source), faster and requiring less training. The procedure can also be modified for high throughput analyses, further reducing the cost and time required for the identification of larval fish and fish eggs when large sample quantities are involved. The discrepancies in identification between the two methods was driven by 3 factors: (1) the lack of interspecific variability in COI for the *Coregonus* genus for DNA barcoding; (2) limited morphological resolution for *Catostomus* and Cyprinidae; and (3) a remaining 21 scattered mismatches affecting a variety of species.

4 Chapter 4: The development of a novel multiplex TaqMan real-time PCR assay for the detection of lake whitefish (*Coregonus clupeaformis*), deepwater sculpin (*Myoxocephalus thompsonii*) and round whitefish (*Prosopium cylindraceum*)

A novel multiplex real-time PCR assay for the detection of lake whitefish (*Coregonus clupeaformis*), deepwater sculpin (*Myoxocephalus thompsonii*) and round whitefish (*Prosopium cylindraceum*) in the Laurentian Great Lakes

(Original Research)

Natalie D.J. Taylor, Andrew M. Zarnke, Christopher M. Somers, Richard G. Manzon,
Joanna Y. Wilson, Douglas R. Boreham

[In review for submission in the Journal of Great Lakes Research]

4.1 Abstract

The accurate identification of larval fish and fish eggs is important for better understanding fish development, spawning grounds, and migration routes. It can also be a critical part of assessing the environmental impacts of industrial processes on fish communities. Current identification methods can be both costly and time consuming. Here we developed a novel multiplex TaqMan real-time PCR assay for the detection and identification of lake whitefish (*Coregonus clupeaformis*), deepwater sculpin (*Myoxocephalus thompsonii*) and round whitefish (*Prosopium cylindraceum*), which are considered Valued Ecological Components (VECs) in the Great Lakes. We used the *Cytochrome b* gene for lake whitefish and the *Cytochrome oxidase I* gene for the other two species to develop species-specific primers and probes for the identification of these species of fish. We were able to differentiate the target VEC species from all other fish tested using cut-off C_t values (16 reference species used). Our assay was validated through randomized blind tests, which confirmed 100% sensitivity and specificity.

Keywords:

Real-time PCR, *Coregonus clupeaformis*, *Myoxocephalus thompsonii*, *Prosopium cylindraceum*

4.2 Introduction

The accurate identification of larval fish and fish eggs is important for many reasons including: acquiring a better understanding of the life cycles of fish, food security (e.g. caviar), and understanding the environmental impact of industrial operations. Historically, larval fish have been identified through morphology, which can be difficult and inaccurate depending on the skill of the taxonomist (8-10). Further, larval fish specimens obtained from environmental studies can be damaged by various processes making their identification much more difficult. The identification of fish eggs is even more problematic, as it relies on the use of light microscopy, which is only accurate to the order, family, or to a group of species that share similar morphologies (11).

The accurate identification of larval fish and fish eggs from the Laurentian Great Lakes is potentially important for understanding the environmental impacts of industries that use large volumes of lake water. For example, electrical generating stations often use large amounts of lake water for once-through cooling systems, which may entrain larval fish and eggs. Entrainment at various power plants has been an important aspect to environmental monitoring for many power plants in North America, to ensure that entrainment is not having a significant impact on the overall biomass equilibrium of various species of fish (74). Lake Huron currently hosts the largest nuclear power plant in the world (Bruce Power) near Tiverton, Ontario (44.338°N, 81.573°W). This generating station operates 8 CANDU reactors that use lake water drawn directly from Lake Huron as part of their steam-condensing system; the two large water intakes (175,000 – 190,000 L/s) periodically draw in small fish and eggs. As part of

understanding the potential environmental impacts of power generation, it is important to know which species of larval fish and fish eggs are entrained in once-through cooling systems.

It is of principal interest to determine whether entrained fish are members of species that have been identified as valuable ecological components (VECs) in the Laurentian Great Lakes (41). Lake whitefish (*Coregonus clupeaformis*) are a VEC because of their economic and cultural value to the commercial fishing industry and First Nations groups (41, 43). This species may spawn in the vicinity of Bruce Power, and is therefore an important species to monitor (41). Deepwater sculpin (*Myoxocephalus thompsonii*), also an identified VEC, are listed as a species of special concern in the Great Lakes, which therefore makes them an important species to monitor (62). These fish have poorly understood spawning activities, but they are known to inhabit the deep parts of cold, well-oxygenated lakes. The relative abundance of deepwater sculpin has been used as an indicator of the status of the deepwater fish community as they are an important prey species for piscivores, such as burbot (*Lota lota*) (45). Round whitefish (*Prosopium cylindraceum*) are thought to be more thermally sensitive than other fish, such as lake whitefish (47, 48). They are a species of special interest for Bruce Power as the once-through cooling system effluent can create a thermal plume in Lake Huron upon discharge (59). Furthermore, round whitefish have been identified as a VEC for both of the other nuclear generating stations in Ontario (Pickering and Darlington; (42). Round whitefish are known to inhabit the Baie du Doré, a bay near Bruce Power, and adults in spawning condition have been regularly captured in the area during the fall (e.g., (59). It is therefore important to be able to accurately identify the larval fish and fish eggs of these important species to allow for a better understanding of the potential impacts of human activity of round whitefish.

An alternative method proposed for the identification of larval fish and fish eggs (in place of morphological identification) is DNA barcoding. DNA barcoding relies on the sequencing of a standard gene, usually the *cytochrome oxidase subunit I* (COI) gene in the mitochondrial genome (19). While DNA barcoding has worked well for the identification of many species, approximately 10% of North America's freshwater fish species do not have adequate interspecific variation in their COI gene for accurate discrimination, including several genera found in the Laurentian Great Lakes (e.g. several species in the *Coregonus* genus – see Ch. 3) (26). Real-Time PCR is an alternative technique that provides rapid identification of target species of fish. Where DNA barcoding requires amplification through PCR which is followed by sequencing, real-time PCR requires only PCR-based amplification for specimen identification, which can be done in real time. Real-time PCR can also circumvent the difficulties associated with a lack of interspecific variation observed in some species when using the COI gene for DNA barcoding. It can do this by targeting alternative genes with higher interspecific variation than is seen in the COI gene. DNA barcoding is currently confined to the COI gene, as comprehensive databases do not exist for the other potential target genes (27). Real-time PCR is therefore more versatile than DNA barcoding in terms of potential target genes.

Here we develop a multiplex TaqMan Real-Time PCR assay that rapidly and specifically identifies lake whitefish, deepwater sculpin and round whitefish. Our overall objective was to create a fast and highly specific assay to identify these three VECs, which are of specific importance for monitoring environmental impacts in Lake Huron. TaqMan real-Time PCR was chosen because it allows for the rapid and sensitive detection of all three species simultaneously using a fluorescence assay. Overdyk, Braid, et al. (75) recently designed a real-time PCR assay for the detection of lake whitefish. However, this assay was singleplex, and it was recommended

that a multiplex assay be developed for the detection of important species. Furthermore, the primers and probe designed by Overdyk, Braid, et al. (75) did not meet our criteria (in terms of interspecific variability – there was not an adequate number of variable base pairs between species). We therefore designed a new primer/probe set for lake whitefish, but on *Cyt b*. *Cyt b* was chosen as the target gene for lake whitefish because it provided an adequate level of variation between lake whitefish and the congeneric species and has been successfully used previously to differentiate animal species. For lake whitefish, the COI gene did not have any sites that met our criteria for potential primer/probe locations (in terms of interspecific variability). The COI gene was targeted for deepwater sculpin and round whitefish because it has been widely used for species identification in DNA barcoding (32, 34, 76, 77). We established cut-off C_t values that were used to objectively determine when the DNA from a specimen was from a target species or not. We were able to develop a novel multiplex TaqMan real-time PCR assay with 100% sensitivity and 100% specificity.

4.3 Materials and Methods

4.3.1 Larval Fish Sample Collection

Larval fish samples were collected from Bruce Power using 500 μ M plankton nets. Each net was 0.72m in diameter and anchored 3m below the water surface in the water intake system for 5-140 minutes (median 36 minutes). Specimen collection occurred both during the day and night from March through December in 2013 and 2014. There were a total of 161 sampling efforts. Individual larval fish and eggs were given unique identification numbers and stored in 95% ethanol until analysis. In total, sampling efforts resulted in the collection of 1740 larval fish and 2831 fish eggs, however, not all of these were analysed.

4.3.2 Reference Collection

DNA for the development of this assay was obtained both from larval fish specimens (collected from Bruce Power) and from adult fish and egg specimens (obtained from various specimen libraries across Canada and the United States). All larval fish specimens were identified using both morphology and DNA barcoding, to ensure accuracy of identifications. Morphological identification of larval fish was performed by an expert ichthyologist, who specialized in identifying larval fish from the Laurentian Great Lakes (Golder Associates). All reference DNA originated from either adult fish tissue samples (from various museum collections) or eggs that had been fertilized *in vitro* by identified parent fish. Species of fish collected, identified and used as references included: bloater (*Coregonus hoyi*), burbot, cisco (*Coregonus artedii*), common carp (*Cyprinus carpio*), deepwater sculpin, emerald shiner (*Notropis atherinoides*), lake whitefish, mottled sculpin, (*Cottus bairdii*), mountain whitefish (*Prosopium williamsoni*), pygmy

whitefish (*Prosopium coulterii*), quillback (*Carpoides cyprinus*), rainbow smelt (*Osmerus mordax*), round goby (*Neogobius melanostomus*), round whitefish, slimy sculpin (*Cottus cognatus*), spoonhead sculpin (*Cottus ricei*), spottail shiner (*Notropis hudsonius*), white sucker (*Catostomus commersonii*), walleye (*Sander vitreus*), and yellow perch (*Perca flavescens*).

4.3.3 DNA Extraction

DNA was extracted from (1) individual larval fish, (2) 2-3 fish eggs that were from the same artificial fertilization effort, and (3) various adult fish muscle tissue and fin clips. When the larval fish was >12mm, only part of the larval fish was used for DNA extraction. Extractions were performed using the Qiagen DNEasy Blood and Tissue Kit (Mississauga, ON) following the guidelines outlined by the manufacturer. Final DNA extraction was performed using MilliQ Biology Grade water, or Nuclease-Free water. DNA concentration and 260:280 values were measured to confirm DNA quality using a NanoDrop 1000 spectrophotometer.

4.3.4 DNA Barcoding

Extracted DNA was diluted to 5ng/μL and a PCR reaction was run using the universal fish primers Fish F1 (5'-TCA ACC AAC CAC AAA GAC ATT GCC AC-3') and FishR1 (5'-TAG ACT TCT GGG TGG CCA AAG AAT CA-3'), which amplified a 658bp region of the *cytochrome oxidase I* (COI) gene (18). Reaction conditions were similar to those used by Ward et al. (2005) and consisted of: 1x PCR buffer, 2.5mM of MgCl₂, 0.1μM of each primer, 0.05mM of each dNTP, 0.3125 units of *Taq* DNA polymerase and 10ng of template DNA in a final volume of 25μL. The thermal cycling regime consisted of 2 minutes at 94°C, 35 cycles of 30 seconds at 94°C, 40 seconds at 52°C, and 1 minute at 72°C; the final extension was 72°C for 10 minutes. Samples were then held at 4°C until amplification of the COI gene was verified by

electrophoresis on a 1% Agarose gel. Successfully amplified DNA was purified using the Qiagen MinElute PCR purification kit (Mississauga, ON) following the guidelines outlined by the manufacturer. Final elution was performed with MilliQ Biology Grade water or Nuclease Free water. Samples were sequenced using Sanger sequencing (University of Calgary). COI sequences were compared to the BLAST database (64) and were considered a match to a species when they matched an individual species at $\geq 98\%$ similarity and had a bit score ≥ 1000 . When there was more than one species, or no species, with $\geq 98\%$ similarity, samples were identified to the top-matching genus. Fish species that do not occur in the study region were not included as potential matches.

4.3.5 *Sequence Selection*

COI sequences were obtained from the Barcode of Life Database (BOLD) and *Cyt b* sequences were obtained from GenBank (29). All sequences from a target species were extracted from BOLD and GenBank (regardless of geographic origin) to avoid possible polymorphisms (21 sequences for lake whitefish, 11 sequences for deepwater sculpin and 54 sequences for round whitefish). Reference sequences were extracted from these databases for all North American congeners, all Lake Huron confamilials and all other non-related species of fish known to occur in Lake Huron. Sequences were aligned in SeaView (63) and compared within species to identify sites with low intraspecific variability. Sites were considered to have low intraspecific variability when there was ≤ 1 base pair difference for potential primer locations and no differences in the base pairs for potential probe locations. These sequences were then compared with sequences from other species within the same genus and family to test for interspecific variability. Species-specific bases were included in primers where possible; however, we maximized the number of species-specific bases in the probes and required the potential probe

locations to have ≥ 3 species-specific bases. Sites with both low intraspecific and high interspecific variability were manually chosen as candidates for primer/probe locations.

4.3.6 Primer and Probe Design and Verification/Validation

Primers and probes were designed using Primer3Plus (78). Specificity of candidate primer and probe sequences were verified using the Basic Local Alignment Search Tool (BLAST) (64). Candidate sequences were also analyzed using OligoAnalyzer 3.1 (79) to estimate the melting temperature, GC content, and to analyze possible secondary structures and hetero-dimers. Melting temperatures for primers were targeted for between 40°C and 50°C to assist with compatibility in multiplex reactions, and between 47°C and 56°C for probes. GC content for all primers and probes was restricted to between 30% and 55%. The melting temperature of any possible hairpin was required to be at least 5°C lower than the melting temperature of the primer/probe. In terms of hetero-dimers, the maximum number of consecutive complementary bases between different primers and probes was 6. Final primer and probe sequences for all target species can be found in Table 4.

Table 4 - The primer and probe sequences for detection of species-specific DNA. Bases underlined with are bases that are variable between the target species and 1 other congeneric species, base highlighted with **G are bases that are variable between the target species and 2 or more other congeneric species.**

Target Species	Target Gene	Primer/Probe	Reporter Moiety/Quencher	Sequence	T _m (°C)	Amplicon Size (bp)
Lake Whitefish	<i>Cyt b</i>	Forward	HEX/Black Hole 2	5'-AGT AGC AGA <u>C</u> AT <u>A</u> CT CAT-3'	46.6	101
		Reverse		5'-AGA TGG <u>T</u> GA AGT AGA TAA C-3'	45.6	
		Probe		5'-TGT AGA A CA CCC C TT CAT <u>T</u> AT <u>C</u> AT-3'	53.7	
Deepwater Sculpin	COI	Forward	Texas Red/Black Hole 2	5'-GAC C TA ACA ATC <u>T</u> T C T-3'	41.2	89
		Reverse		5'-GGT T TC ATG TTA <u>A</u> TG <u>A</u> T-3'	40.8	
		Probe		5'-C C T TAC ATC TAG CAG GAA TC-3'	49.7	
Round Whitefish	COI	Forward	FAM/Black Hole 1	5'-A G T ATC AAA C A C CCC TTT-3'	48.1	170
		Reverse		5'- <u>T</u> AG GTG CTG ATA CAG AAT-3'	46.9	
		Probe		5'-C A G GTA TTA CAA T A C <u>T</u> GC TTA CG-3'	51.1	

FAM, 6-carboxyfluorescein; HEX, Hexachlorofluorescein; T_m, melting temperature; bp, base pairs

The functionality of primer sets was verified by running a PCR reaction followed by electrophoresis of the product on a 1% Agarose gel to verify amplification. PCR reactions were prepared on ice and occurred in 25 µL volumes containing 1X iQ Supermix (Bio-Rad Laboratories, Mississauga, ON), 12.5 pmol of each primer and 157.5 ng of DNA. Reaction conditions consisted of 95°C for 2 minutes, followed by 40 cycles of 95°C for 11 seconds, 43°C for 30 seconds and 72°C for 10 seconds. Primers were considered to be functional when there was a strong band present for the target species on the agarose gel (expected amplicon size in Table 4). Primers were also verified by running them on an annealing temperature gradient (between 34°C and 50°C) to ensure efficient annealing at multiplex temperatures. Probes were

then verified using a Chromo 4 Real-Time PCR thermocycler (Bio-Rad Laboratories, Mississauga, ON) with the same reaction conditions previously specified, and with the addition of 7.5 pmol of each probe. Primer and probe sets were considered to be functional when there was amplification of the target DNA within 30 cycles. The amount of template DNA was kept consistent to ensure appropriate C_t cut-off values, as the C_t is dependent on the initial amount of template DNA.

Once the functionality of each primer/probe set was verified and amplification of the target species DNA was confirmed, the primer/probe sets were tested for specificity with DNA from 17 non-target reference species. These reference species were chosen either because they were closely related to the target species (same genus), or they were species of fish that co-occurred in Lake Huron. When there was no non-specific amplification, reference specimen DNA was run in a PCR reaction containing fish primers (Fish F1 and Fish R1) to ensure that the DNA was amplifiable (18). Primer and probe sets were also tested against unrelated fish species that are known to occur in Lake Huron. Once each individual primer/probe set was verified, they were combined in a multiplex reaction. The multiplex reaction was validated using a minimum of 3 replicates of each target species DNA and a minimum of 3 replicates of each reference species DNA. A 10-fold serial dilution was used to generate a standard curve (initial amount of DNA was 160ng) for each target species. The validation multiplex was run 3 separate times with different DNA from different specimens each time.

Using the results from these multiplex reactions, we established cut-off C_t values for each target species. Cut-off C_t values were needed to establish a definitive point when observed amplification can be considered to be from a target species. Cut-offs were based on the 95% confidence intervals (CI) around the average C_t of the target species and the next-lowest average

C_t of a reference species. Cut-off values were placed at the mid-point between the upper 95% CI of the target species and the lower 95% CI of the reference species. To ensure specificity, the validated multiplex reaction was also run under conditions where the analyst was blind to the previously established identity of each specimen.

4.4 Results

4.4.1 Specificity

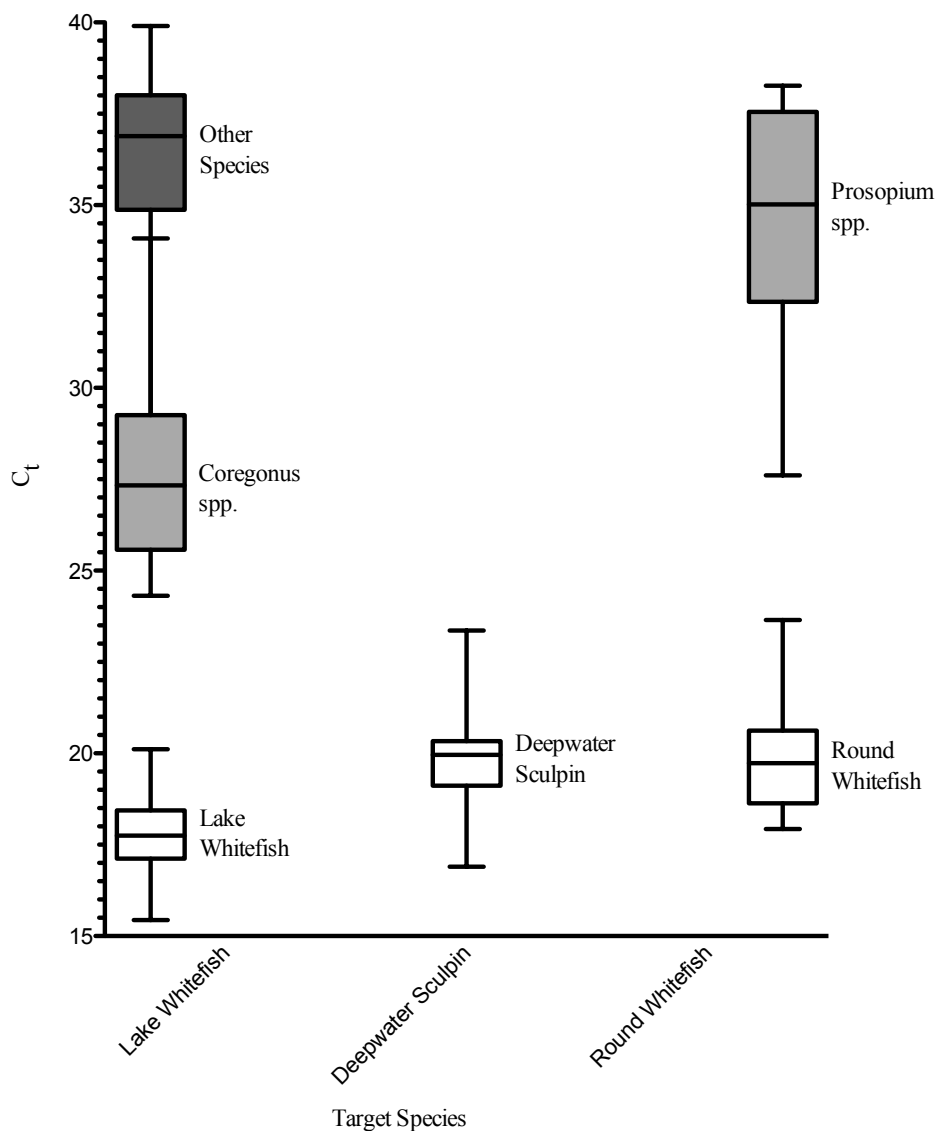
The average C_t for all lake whitefish multiplex assays run was 17.9 (95% Confidence interval (CI): 16.9 – 18.8) (Table 5). To ensure that the assay could differentiate lake whitefish from other members of the *Coregonus* genus, both bloater (*Coregonus hoyi*) and cisco (*Coregonus artedii*) were used as reference specimens. The average C_t for *Coregonus* fish (excluding lake whitefish) was 28.0 (95% CI: 25.2 – 30.8) (Figure 3). The cut-off C_t for lake whitefish was therefore determined to be at a C_t of 22, as this was the approximate median between the upper 95% CI of lake whitefish and the lower 95% CI of the other *Coregonus* species. With this cut-off value, we were able to differentiate lake whitefish from other members of the *Coregonus* genus in our blind and multiplex validation tests 100% of the time with no false positives. There were no instances where the C_t for lake whitefish DNA was above our cut-off value of 22 (maximum C_t was 20.11); we therefore had no false negatives. There was also some amplification with other non-*Coregonus* species of fish (e.g. round whitefish). However, the C_t values for the other species of fish were always greater than those of the members of the *Coregonus* genus (maximum C_t for *Coregonus* specimens was 34.09, minimum C_t for non-*Coregonus* specimens was 35.98).

Table 5 - The average C_t of target DNA and congeneric DNA using the designed primers and probes. Maximum and minimum C_t values are the maximum and minimum over all runs and did not necessarily occur in the same test run. SE denotes standard error and N denotes the number of samples.

Primer/Probe Target	Relative species	Average Target C_t (\pm SE), N	Average C_t of Relative Species (\pm SE), N	Maximum C_t of Target DNA	Minimum C_t of Non-Target DNA	Cut-off C_t
Lake Whitefish	Bloater and Cisco	17.9 (\pm 0.4), 31	28.0 (\pm 0.9), 28	20.1	25.6	22
Deepwater Sculpin	Slimy, Mottled and Spoonhead Sculpin	19.7 (\pm 0.3), 21	No C_t s detected, 27	23.4	No C_t s detected	25
Round Whitefish	Mountain and Pygmy Whitefish	19.77 (\pm 0.3), 31	34.5 (\pm 1.6), 33	21.3	27.6	25

The average C_t for deepwater sculpin was 19.88 (95% CI: 19.1 – 20.6) (Table 5). We ran the assay with DNA from the related species mottled sculpin (*Cottus bairdii*), slimy sculpin (*Cottus cognatus*) and spoonhead sculpin (*Cottus ricei*); our assay was able to differentiate deepwater sculpin from congeners, as well as all other species of fish that were tested (Figure 3). There was no non-specific amplification detected with non-target DNA when analyzing the deepwater sculpin results. The highest C_t observed with the reference deepwater sculpin was 23.4; we were therefore established a conservative cut-off C_t value of 25 for this species. Using the cut-off C_t , our assay was able to discriminate deepwater sculpin with 100% sensitivity and 100% specificity.

Figure 3 – Box and whisker plot showing the mean, quartiles and maximum and minimum C_t values for all target and non-target species. Target species are in white, closely related species are in light grey, and other species are in dark grey.



The average C_t for round whitefish was 19.7 (95% CI: 18.9 – 20.4) (Table 5). To ensure that the assay was specific, we tested it with DNA from both pygmy and mountain whitefish. There was some amplification with the other *Prosopium* species; however, the average C_t was 34.53 (95% CI: 30.6 – 38.5) (Figure 3). We therefore established a conservative cut-off C_t value for round

whitefish of 25. With this cut-off C_t our assay was 100% sensitive and 100% specific for round whitefish.

4.4.2 Application of Assay for Larval Fish and Fish Egg Identification

In order to test the accuracy of the assay for use in the identification of larval fish, we performed a blind test, where the technician was unaware of the identity of the larval fish (e.g. Figure 4). The test included fish DNA from 12 different species, including close relatives of the target DNA. We were able to accurately identify the target specimens 100% of the time when using the established cut-off C_t values described above.

4.4.3 Cost of Real-Time PCR

An important consideration for industries and other scientists is the cost of an assay compared with the cost of an alternative. We calculated the cost of running our real-time PCR per fish sample to be \$3.98 CAD (Table 6). This included the cost of all the primers and probes, the DNA extraction, other reagents (i.e. nuclease-free water) and labour. The cost of morphological identification by the expert ichthyologist was \$13.30 CAD per specimen and the cost in our lab for DNA barcoding was \$12.68 CAD per specimen. However, it is important to note that while morphological identification and DNA barcoding can identify specimens, real-time PCR can only determine if the target DNA was present or absent in the sample. Real-time PCR was significantly less expensive than DNA barcoding because it does not require many of the steps that DNA barcoding requires (e.g. sequencing).

Figure 4 - Example of a blind multiplex run where identifications were assigned by a blind technician and then confirmed with the identifications from DNA barcoding. Individual data points represent individual specimens.

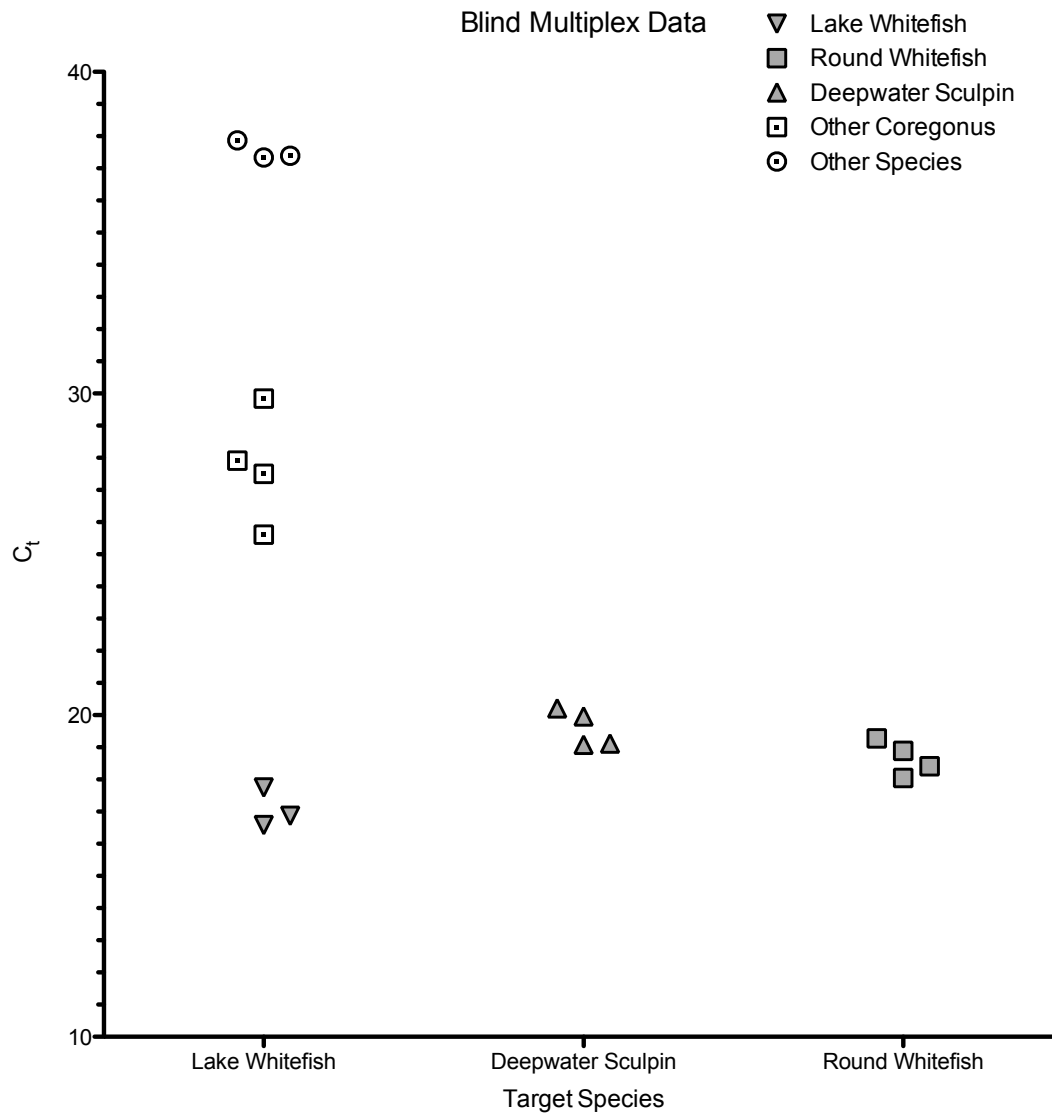


Table 6 - Cost breakdown for each component of the multiplex TaqMan real-time PCR. Cost per specimen was calculated based on the number of specimens that could be analyzed with each component.

Item	Total Cost (CAD)	Number of Specimens that can be Analyzed	Cost per Specimen (CAD)
Qiagen DNEasy Blood and Tissue Kit	\$575.00	250	\$2.30
Bio Rad iQ Super Mix	\$752.00	1000	\$0.752
Lake Whitefish F Primer	\$9.70	5,858	\$0.0017
Lake Whitefish R Primer	\$10.10	5,666	\$0.0018
Lake Whitefish Probe	\$315.00	2,466	\$0.13
Deepwater Sculpin F Primer	\$8.90	7,720	\$0.0012
Deepwater Sculpin R Primer	\$9.30	6,505	\$0.0014
Deepwater Sculpin Probe	\$535.00	1,733	\$0.31
Round Whitefish F Primer	\$10.00	2,573	\$0.0038
Round Whitefish R Primer	\$10.00	2,157	\$0.0046
Round Whitefish Probe	\$414.13	2,586	\$0.16
Nuclease-Free Water	\$40.24	20,000	\$0.002
Labour	\$500	1,612	\$0.31
Total			\$3.98

4.5 Discussion

In this study, we designed a multiplex TaqMan real-time PCR assay for the presence/absence of lake whitefish, deepwater sculpin and round whitefish. This assay can be applied to both adult fish (for forensic applications) as well as fish in the early stages of development, which are the most difficult to identify (8, 9). It was important that this assay be able to differentiate lake whitefish from other members of the *Coregonus* genus, which are notoriously difficult to differentiate (26, 80, 81). When lake whitefish was the target, there was cross amplification with bloater (*Coregonus hoyi*) and cisco (*Coregonus artedii*) because of high genetic similarity within the *Coregonus* genus. Initially, we targeted the COI gene for lake whitefish identification; however, there were no locations on the COI gene that contained sufficient interspecific variability to differentiate lake whitefish from other *Coregonus* species. We therefore targeted the *Cyt b* gene, which has been used for species differentiation in other real-time PCR assays (32, 34, 76). We were able to find a location with low intraspecific variability and high interspecific variability in the *Cyt b* gene. There were 5 locations on the gene where lake whitefish was different from the other *Coregonus* species and these locations were incorporated into the forward primer and the probe (2 in the forward primer and 3 in the probe). It should be noted that there were no differences in this region between lake whitefish and *Coregonus baunti*; however, *Coregonus baunti* is only found in Siberia, and therefore does not have an overlapping distribution with lake whitefish. Thus, our real-time PCR approach is an efficient tool for identification of lake whitefish early life stages in the Laurentian Great Lakes, and likely elsewhere in North America.

We were able to differentiate deepwater sculpin from other sculpin species found in the Laurentian Great Lakes. This is because of the high number of differences in the COI gene that we were able to incorporate into the primers and probe (6-7 differences per species). However, we predict that our assay would not be able to differentiate deepwater sculpin from fourhorn sculpin (*Myoxocephalus quadricornis*) because of the similarity in their genetic sequences. There was only 1 base pair difference between deepwater sculpin and fourhorn sculpin that could be incorporated into the primers and probe (this difference was incorporated into the reverse primer). Deepwater sculpin and fourhorn sculpin are closely related species; in fact, there has been debate about whether deepwater sculpin is a subspecies of fourhorn sculpin (82). However, fourhorn sculpin and deepwater sculpin do not have overlapping distributions (83, 84). Therefore, if the geographical context is taken into consideration when identifying these fish, there should not be any false positives.

Finally, our assay was able to effectively differentiate round whitefish from all other species of fish. While mountain whitefish does not occur in the Laurentian Great Lakes, pygmy whitefish does occur in Lake Superior (85). Our assay was able to differentiate round whitefish from these closely related fish because of the high level of interspecific variability at the primer/probe locations between round whitefish and mountain/pygmy whitefish (total of 5 and 6 consistent base pair differences respectively).

Another important part of this study was the identification of cut-off C_t values for each target species of fish. Many other studies have employed the use of cut-off values as an objective way of separating positive and negative results (86-88). Our cut-offs were determined by comparing the CIs of the target species with the CIs of the species with the next lowest C_t . The cut-off was chosen so as to balance the likelihood of false positives with the likelihood of false negatives.

These cut-off values are specific to the amount of template DNA that was used in the reaction. If a different amount of DNA is used then a new cut-off value will need to be determined as the C_t is partially dependent of the amount of template DNA present, amongst other factors (e.g. the specificity of the primers and probe). Therefore, should a different amount of template DNA be used than those reported in this study, new cut-off C_t values should be determined.

A potential advantage of real-time PCR is the possibility to estimate the amount of template DNA. For quantitative real-time PCR, it is important to have efficiencies that are between 90-105% (89). Our multiplex reaction did not result in an assay where the efficiency was consistently between 90 and 105% as the reaction conditions were not optimal for all primer/probe sets. For this reason, we do not recommend these primers and probes for use in quantitative applications of real-time PCR. However, the purpose of our design of this real-time PCR assay was not to quantify the amount of template DNA, but rather to differentiate species based on their DNA.

A major advantage of real-time PCR is that it can be adapted for field-based identifications (90). The rapid identification of ichthyoplankton in the field would allow researchers to conduct their research more quickly and efficiently. Rather than sampling multiple sites and hoping that one would contain the target species, real-time PCR in the field would allow the researcher to be sure of which sites were of interest (e.g. contained the species of interest), and would therefore allow them to be more focused in their efforts. Furthermore, the use of a TaqMan probe (as opposed to an intercalating dye such as SYBR Green) adds an additional point of specificity for species differentiation. By using an internal probe, we were able to add a minimum of 3 additional species-specific bases for each target species of fish. This allowed us to accurately discriminate our target species from other species of fish. Finally, real-time PCR also eliminates the need for

post-PCR processing, something that is still required in DNA barcoding. This makes real-time PCR faster and cheaper than other PCR-based methods that require post-PCR processing.

Specifically analyzing the cost of real-time PCR, we determined that it is a cost-effective option for the identification of larval fish and fish eggs when looking for certain species. Our multiplex reaction costs \$3.98 CAD per fish, which is much lower than the \$13.30/fish that it cost our lab to identify them morphologically, or the \$12.68/fish that it cost our lab for DNA barcoding identification. The most expensive part of real-time PCR is the development of the assay, as it may take multiple primers and probes to find the combination that is the most effective, however, once the assay has been developed it is a low-cost option that can be modified for high-throughput applications.

Our assay is well suited for the identification of larval fish and fish eggs in environmental assessments for major industries on the Laurentian Great Lakes. It is important to identify ichthyoplankton in environmental studies for several reasons, including: having a better understanding of the scale of the environmental impact of human activities and how reparations may be made, and managing the populations that are impacted by our activities. Our assay is well suited for these applications because it can rapidly determine if specimens are VECs and in what abundance compared to other species of fish. In addition, it is also both low-cost and a good alternative to morphological identification, as well as other molecular identification methods for the identification of lake whitefish, deepwater sculpin and round whitefish. The use of real-time PCR will allow for the rapid identification of these species and provide industries on the Laurentian Great Lakes with a more accurate picture of their environmental impact. Real-time PCR has a wide breadth of application possibilities. Here, we report another novel

application for multiplex real-time PCR for use in the identification of larval fish and fish eggs in environmental samples.

4.6 Acknowledgements

The researchers would like to thank the following people and institutions for providing reference tissue samples: Dr. James H. Johnson from the Great Lakes Science Centre, Dr. David Marcogliese, Dr. Doug Watkinson, Dr. Andrew Simons from the James Ford Bell Museum of Natural History, and Dr. Dirk Steinke from the Biodiversity Institute of Ontario. This research was funded in part by Bruce Power, Mitacs and NSERC.

5 Chapter 5: Conclusions and Future Directions

In recent years, molecular techniques have been growing in popularity for the identification of various types of species (91-93). Prior to the use of DNA and proteins for species identification, researchers relied solely on morphological techniques for their identifications. Morphological identification of larval fish and fish eggs is dependent on the gross physical integrity of the specimens and the expertise of the taxonomist who is responsible for their identifications. These subjective identifications can result in varying identifications of the same larval fish specimens by different taxonomists (10). Molecular techniques were developed, in part, to provide a more objective method for species identification and to provide identifications for specimens that were no longer grossly intact (e.g. for canned meat product investigations) (17). Even with increased application of these novel techniques, no studies have previously compared morphological identifications of larval Canadian freshwater fishes with molecular identifications of the same specimens.

DNA has been used in many studies to identify different species of fish (13, 16, 54, 72, 94). The COI gene is widely employed for DNA barcoding and there are comprehensive databases that exist for the comparison of COI genetic sequences (27-29). We therefore chose DNA barcoding as the alternative molecular identification technique for comparison with morphological identifications. The first objective of this project was to test the hypothesis that DNA barcoding would be able to provide a greater species resolution than morphological identifications. The resolution of morphological identifications is limited as a result of the subjective nature of morphological identifications and the difficulty of identifying larval fish specimens. Secondly, we verified that DNA barcoding could identify fish eggs, as fish eggs are difficult to identify

morphologically and in general can only be identified to the order or family using morphological identification (11).

Contrary to our hypothesis, we discovered that the ability of morphological identifications to resolve larval fish specimens to the species is not significantly different than the ability of DNA barcoding to resolve ichthyoplankton specimens. The few differences that were seen between the two techniques were caused by three main reasons: (1) the inability of DNA barcoding to differentiate bloater, cisco and kiyi – these specimens were therefore only identified to the *Coregonus* genus; (2) the inability of morphological identifications to resolve specimens from the *Catostomus* genus beyond the genus; and (3) the inability of morphological identifications to identify specimens from the *Cyprinidae* family beyond the family level. We also found that DNA barcoding could identify approximately 50% of the analyzed fish eggs to the species. We could not compare the morphological identification of eggs with DNA barcoding because the eggs were not identified morphologically.

We therefore concluded that despite the inability of DNA barcoding to resolve a specific subset of species (coregonids), it is still a more robust means of species identification than morphological identification for other taxonomic groups (26). Overall, we found that although both approaches are not able to completely resolve all species, DNA barcoding has advantages, including the ability to be modified for high throughput scenarios, it functions well regardless of the gross physical integrity of the specimen and it is more cost-effective than morphological identification when graduate students are used as the labour source. We would, however, suggest that for future use, another gene be used in addition to the COI gene for DNA barcoding. With the development of next-generation sequencing (NGS) and the various sequencing platforms that are being developed, it is possible to sequence multiple genes at once, and this could allow for

multiple loci to be used for identification verification (95). Primers for both the COI gene and another target gene (e.g. *Cyt b*) could be run in a multiplex reaction and sequenced simultaneously so as to not add a significant increase in cost (96, 97). The only barrier to this would be the development of a comprehensive database for the alternative gene. However, databases have already been developed for the COI gene, which demonstrates that it is possible to rapidly assemble a comprehensive database for DNA barcoding.

While DNA barcoding has many benefits as a technique to identify most species of Canadian freshwater fishes, it is still limited by a relatively high cost per sample as is unable to resolve specific species in multiple genera. We were particularly interested in identifying VECs as identified by Bruce Power (41). TaqMan real-time PCR is a cost effective method for the rapid identification of targeted species simultaneously in a multiplex reaction. We therefore developed a novel multiplex TaqMan real-time PCR assay that was able to detect lake whitefish, deepwater sculpin and round whitefish. We utilized cut-off C_t values to discriminate our target species from their closely related congeneric species. When tested in both known and blind multiplex reactions, our assay (with our cut-off values) was able to detect 100% of the target DNA, resulting in no false negatives. It was also able to discriminate between the target DNA and DNA from all other species, which resulted in no false positives. Overall, we were able to design and validate a novel multiplex TaqMan real-time PCR assay that was both sensitive and specific for the three target species, supporting our original hypothesis.

The scope of a study is directly impacted by the cost of identifications, thus we determined the exact cost of each form of fish identification approach on a per specimen basis. We found that morphological identification was the most expensive as it required the skills of a highly trained ichthyologist with years of training who specialized in the identification of larval fish from the

Laurentian Great Lakes. DNA barcoding was less expensive than morphological identification and had the potential for the cost to be reduced further by streamlining the process that was used (e.g. by doing in-house sequencing of the specimens); however, it was still more than \$10 CAD per larval fish. The cost could potentially be greatly reduced through the use of NGS. NGS would allow for the simultaneous sequencing of hundreds of specimens, thereby reducing the sequencing cost per specimen (98). The multiplex TaqMan real-time PCR assay was expensive to develop, but after validation it cost less than \$4 CAD for the differentiation of each fish, with the majority of the cost stemming from the DNA extraction process. The multiplex assay also has the potential to be used in high throughput applications, which could further reduce the cost. Therefore, when trying to identify a few species (e.g. VECs), TaqMan real-time PCR is the most cost-effective option. However, if the goal is to identify all ichthyoplankton specimens, DNA barcoding is the more appropriate option because of its versatility (compared with real-time PCR), and its advantages over morphological identification, namely in its ability to identify specimens that had sustained gross damage, lower cost and the decreased amount of training required.

In the future, in addition to the VECs identified here, there are many other VECs that need to be rapidly identified by industries on the Laurentian Great Lakes (e.g. smallmouth bass [*Micropterus dolomieu*] and brook trout [*Salvelinus fontinalis*]) (41). Multiplex assays should be developed for their detection as well, which could be used concurrently with the assay developed here. This technology has a wide variety of applications that extend beyond the Laurentian Great Lakes. DNA barcoding can be used for the identification of any species of animal, as long as there is a comprehensive database that exists and there is an adequate “barcoding gap” (19, 27). For real-time PCR, the primers and probes presented here can be used concurrently in multiplex

format, or individually, further expanding the geographic area in which they are useful for species differentiation. Overall, this project demonstrated the utility and versatility of molecular methods for species identification.

References

1. Vacchi M, La Mesa M, Dalu M, Macdonald J. Early life stages in the life cycle of Antarctic silverfish, *Pleuragramma antarcticum* in Terra Nova Bay, Ross Sea. *Antarct Sci.* 2004;16(3):299-305.
2. Connell AD. A 21-year ichthyoplankton collection confirms sardine spawning in KwaZulu-Natal waters. *Afr J Mar Sci.* 2010;32(2):331-6.
3. Brochier T, Ramzi A, Lett C, Machu E, Berraho A, Freon P, et al. Modelling sardine and anchovy ichthyoplankton transport in the Canary Current System. *J Plankton Res.* 2008;30(10):1133-46.
4. López-Sanz À, Stelzenmüller V, Maynou F, Sabatés A. The influence of environmental characteristics on fish larvae spatial patterns related to a marine protected area: The Medes islands (NW Mediterranean). *Estuar Coast Shelf Sci.* 2011;92(4):521-33.
5. Ahlstrom EH, Moser HG. Eggs and larvae of fishes and their role in the systematic investigations and in fisheries. *Rev Trav Inst Peches Marit.* 1976;40(3 and 4):379-98.
6. Thome C, Mitz C, Somers CM, Manzon RG, Boreham DR, Wilson JY. Incubation of Lake Whitefish (*Coregonus clupeaformis*) embryos in cooling water discharge and the impacts of fluctuating thermal regimes on development. *Can J Fish Aquat Sci.* 2016;0(0):1-35.
7. Rago PJ. Production foregone: an alternative method for assessing the consequences of fish entrainment and impingement losses at power plants and other water intakes. *Ecol Modell.* 1984;24:79-111.
8. Strauss RE, Bond CE. Taxonomic Methods: Morphology. In: Schreck C, Moyle P, editors. *Methods for Fish Biology*. Maryland: American Fisheries Society; 1990. p. 109-40.
9. Teletchea F. Molecular identification methods of fish species- reassessment and possible applications. *Rev Fish Biol Fish.* 2009;19:265-93.
10. Ko HL, Wang YT, Chiu TS, Lee MA, Leu MY, Chang KZ, et al. Evaluating the accuracy of morphological identification of larval fishes by applying DNA barcoding. *PLoS One.* 2013;8(1):e53451.
11. Shao K-T, Chen K-C, Wu J-H. Identification of marling fish eggs in Taiwan using light microscopy, scanning electric microscopy and mtDNA sequencing. *CSIRO.* 2002;53:355-65.
12. Hyde JR, Lynn E, Humphreys R, Musyl M, West AP, Vetter R. Shipboard identification of fish eggs and larvae by multiplex PCR, and description of fertilized eggs of blue marlin, shortbill spearfish, and wahoo. *Mar Ecol Prog Ser.* 2005;286:269-77.
13. Ward RD, Hanner R, Hebert PD. The campaign to DNA barcode all fishes. *J Fish Biol.* 2009;74(2):329-56.
14. Ward RD, Grewe PM. Appraisal of molecular genetic techniques in fisheries. *Rev Fish Biol Fish.* 1994;4:300-25.
15. Chow S, Nohara K, Tanabe T, Itoh T, Tsujii S, Nishikawa Y, et al. Genetic and morphological identification of larval and small juvenile tunas (Pisces: Scombridae) caught by a mid-water trawl in the western Pacific. *Bulletin: Fisheries Research Agency.* 2003;8:1-14.

16. Hubert N, Delrieu-Trottin E, Irisson JO, Meyer C, Planes S. Identifying coral reef fish larvae through DNA barcoding: a test case with the families Acanthuridae and Holocentridae. *Mol Phylogenet Evol.* 2010;55(3):1195-203.
17. Hebert PD, Ratnasingham S, deWaard JR. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc Biol Sci.* 2003;270 Suppl 1:S96-9.
18. Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PD. DNA barcoding Australia's fish species. *Philos Trans R Soc Lond B Biol Sci.* 2005;360(1462):1847-57.
19. Hebert PD, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc Biol Sci.* 2003;270(1512):313-21.
20. Barrett RDH, Hebert PDN. Identifying spiders through DNA barcodes. *Can J Zool.* 2005;83:481-91.
21. Moritz C, Cicero C. DNA barcoding: promise and pitfalls. *PLoS Biol.* 2004;2(10):e354.
22. Avise JC. *Phylogeography: The history and formation of species.* Cambridge Mass.: Harvard University Press; 2000.
23. Galtier N, Nabholz B, Glemin S, Hurst GD. Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol.* 2009;18(22):4541-50.
24. Gissi C, Iannelli F, Pesole G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. *Heredity (Edinb).* 2008;101(4):301-20.
25. Rubinoff D, Cameron S, Will K. A genomic perspective on the shortcomings of Mitochondrial DNA for "barcoding" identification. *J Hered.* 2006;96(6):581-94.
26. April J, Mayden RL, Hanner RH, Bernatchez P. Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Science.* 2011;108(26):10602-7.
27. Meyer CP, Paulay G. DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.* 2005;3(12):e422.
28. Barcode of Life Database (BOLD) [Internet]. 2016. Available from: <http://www.boldsystems.org/index.php/>.
29. GenBank: NCBI; 2016 [Available from: <http://www.ncbi.nlm.nih.gov/genbank/>].
30. Ponchel F, Toomes C, Bransfiels K, Leong FT, Douglas SH, Field SL, et al. Real-time PCR based on SYBR-Green I fluorescence: An alternative to the TaqMan assay for a real-time quantification of gene rearrangements, gene amplifications and micro gene deletions. *BCM Biotechnology.* 2003;3(18):1-13.
31. Fajardo V, González I, Rojas M, García T, Martín R. A review of current PCR-based methodologies for the authentication of meats from game animal species. *Trends Food Sci Technol.* 2010;21(8):408-21.
32. You J, Huang L, Zhuang J, Mou Z. Species-specific multiplex real-time PCR assay for identification of deer and common domestic animals. *Food Sci Biotechnol.* 2013;23(1):133-9.

33. Trotta M, Schonhuth S, Pepe T, Cortesi ML, Puyet A, Bautista JM. Multiplex PCR method for use in real-time PCR for identification of fish fillets from grouper (*Epinephelus* and *Mycteroperca* species) and common substitute species. J Agric Food Chem. 2005;53:2039-45.
34. Dooley JJ, Paine KE, Garrett SD, Brown HM. Detection of meat species using TaqMan real-time PCR assays. Meat Sci. 2004;68(3):431-8.
35. Kesmen Z, Gulluce A, Sahin F, Yetim H. Identification of meat species by TaqMan-based real-time PCR assay. Meat Sci. 2009;82(4):444-9.
36. Rougemont M, Van Saanen M, Sahli R, Hinrikson HP, Bille J, Jatton K. Detection of four Plasmodium species in blood from humans by 18S rRNA gene subunit-based and species-specific real-time PCR assays. J Clin Microbiol. 2004;42(12):5636-43.
37. Lehmann LE, Hunfeld KP, Emrich T, Habershausen G, Wissing H, Hoeft A, et al. A multiplex real-time PCR assay for rapid detection and differentiation of 25 bacterial and fungal pathogens from whole blood samples. Med Microbiol Immunol. 2008;197(3):313-24.
38. Dalmaso A, Fontanella E, Piatti P, Civera T, Rosati S, Bottero MT. A multiplex PCR assay for the identification of animal species in feedstuffs. Mol Cell Probes. 2004;18(2):81-7.
39. Lopez R, Asensio C, Guzman MM, Boonham N. Development of real-time and conventional RT-PCR assays for the detection of potato yellow vein virus (PYVV). J Virol Methods. 2006;136(1-2):24-9.
40. Mackay IM. Real-time PCR in the microbiology laboratory. Clin Microbiol Infect. 2004;10:190-212.
41. O'Neill JE. Environmental assessment study report. Bruce Power; 2005.
42. Ecomatrix. Literature review and field investigations of round whitefish habitat along the north shore of Lake Ontario. Ecomatrix Incorporated; 2014.
43. Luna SM, Bailly N. Coregonus clupeaformis (Mitchill, 1818); Lake Whitefish 2016 [Available from: <http://www.fishbase.org/summary/Coregonus-clupeaformis.html>].
44. Freshwater Landings: Government of Canada; 2016 [Available from: <http://www.dfo-mpo.gc.ca/stats/commercial/land-debarq/freshwater-eaudouce/2014-eng.htm>].
45. COSEWIC. COSEWIC Assessment and update status report on the deepwater sculpin *Myoxocephalus thompsonii* (Western and Great Lakes-Western St. Lawrence populations) in Canada. Ottawa; 2006.
46. Deepwater Sculpin Great Lakes Populations: Government of Canada; 2006 [Available from: http://www.registrelep-sararegistry.gc.ca/species/speciesDetails_e.cfm?sid=76].
47. Dunford WE. Whitefish Spawning Study. Ontario Hydro Research Division; 1980.
48. Edsall TA, Rottiers DV. Temperature tolerance of young-of-the-year lake whitefish, *Coregonus clupeaformis*. Journal of the Fisheries Research Board of Canada. 1976;33:177-80.
49. Moura T, Silvia M, Figueiredo I, Neves A, Muñoz P, Coelho M, et al. Molecular barcoding of north-east Atlantic deep-water sharks: species identification and application to fisheries management and conservation. Mar Freshw Res. 2008;59:214-23.

50. Valdez-Moreno M, Vásquez-Yeomans L, Elías-Gutiérrez M, Ivanova NV, Hebert PDN. Using DNA barcodes to connect adults and early life stages of marine fishes from the Yucatan Peninsula, Mexico: potential in fisheries management. *Mar Freshw Res.* 2010;61(6):655-71.
51. Richardson DE, Vanwyke JD, Exum AM, Cowen RK, Crawford DL. High-throughput species identification: from DNA isolation to bioinformatics. *Mol Ecol Notes.* 2007;7(2):199-207.
52. Clare EL, Lim BK, Engstrom MD, Eger JL, Hebert PDN. DNA barcoding of Neotropical bats: species identification and discovery within Guyana. *Mol Ecol Notes.* 2007;7(2):184-90.
53. Ratnasingham S, Hebert PD. A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS One.* 2013;8(7):e66213.
54. Hubert N, Hanner R, Holm E, Mandrak NE, Taylor E, Burrige M, et al. Identifying Canadian freshwater fishes through DNA barcodes. *PLoS One.* 2008;3(6):e2490.
55. Hebert PD, Gregory TR. The promise of DNA barcoding for taxonomy. *Syst Biol.* 2005;54(5):852-9.
56. Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN. Universal primer cocktails for fish DNA barcoding. *Mol Ecol Notes.* 2007;7(4):544-8.
57. Caterino MS, Tishechkin AK. DNA identification and morphological description of the first confirmed larvae of Heterocerinae (Coleoptera: Histeridae). *Syst Entomol.* 2006;31(3):405-18.
58. Pegg GG, Sinclair B, Brinskey L, Aspden WJ. MtDNA barcode identification of fish larvae in the southern Great Barrier Reef, Australia. *Sci Mar.* 2006;70S2(October):7-12.
59. Graham CF, Eberts RL, Morgan TD, Boreham DR, Lance SL, Manzon RG, et al. Fine-scale ecological and genetic population structure of two whitefish (coregoninae) species in the vicinity of industrial thermal emissions. PDF. *PLoS One.* 2016;11(1):e0146656.
60. Overdyk LM, Braid HE, Crawford SS, Hanner RH. Extending DNA barcoding coverage for Lake Whitefish (*Coregonus clupeaformis*) across the three major basins of Lake Huron. *De Gruyter open.* 2015;3:59-65.
61. Schlei OL, Crete-Lafreniere A, Whiteley AR, Brown RJ, Olsen JB, Bernatchez L, et al. DNA barcoding of eight North American coregonine species. *Mol Ecol Resour.* 2008;8(6):1212-8.
62. Committee on the status of endangered wildlife in Canada [Internet]. Government of Canada. 2015 [cited 2016].
63. Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 2010;27(2):221-4.
64. Nucleotide BLAST [Internet]. NCBI. 2016. Available from: <http://blast.ncbi.nlm.nih.gov/Blast.cgi>.
65. Blood DM, Matarese AC. Larval development and identification of the genus triglops. NOAA (National Oceanic and Atmospheric Administration) Professional Paper NMFS. 2010:1-50.

66. Richardson SL, Washington BB. Guide to the identification of some sculpin (cottidae) larvae from marine and brackish waters off Oregon and adjacent areas in the northeast Pacific. NOAA (Natl Oceanic Atmos Adm) Tech Rep NMFS (Natl Mar Fish Serv) Circ. 1980;430:1-61.
67. Khan NY, Faber DJ. A comparison of the larvae of the deepwater and fourhorn sculpin, *Myoxocephalus quadricornis* L. from North America. I. Morphological development. The Early Life History of Fish. 1974:703-12.
68. Fuiman LA, Conner JV, Lathrop BF, Buynak GL, Snyder DE, Loos JJ. State of the Art of Identification for Cyprinid Fish Larvae from Eastern North America. Trans Am Fish Soc. 1983;112(2B):319-32.
69. Snyder DE. Catostomid larvae and early juveniles of the Rio Grande basin in Colorado. Colorado Division of Wildlife; 1998.
70. Kay LK, Wallus R, Yeager BL. Reproductive biology and Early life history of fishes in the Ohio River Drainage. Chattanooga, Tennessee: Tennessee Valley Authority; 1994.
71. Fuiman LA, Witman DC. Descriptions and Comparisons of Catostomid Fish Larvae: *Catostomus catostomus* and *Moxostoma erythrurum*. Trans Am Fish Soc. 1979;108(6):604-19.
72. Kochzius M, Seidel C, Antoniou A, Botla SK, Campo D, Cariani A, et al. Identifying Fishes through DNA Barcodes and Microarrays. PLoS One. 2010;5(9):e12620.
73. Sajdak SL, Phillips RB. Phylogenetic relationships among *Coregonus* species inferred from the DNA sequence of the first internal transcribed spacer (ITS1) of ribosomal DNA. Can J Fish Aquat Sci. 1995;54:1494-503.
74. Barnthouse LW. Impacts of entrainment and impingement on fish populations: A review of the scientific evidence. Environ Sci Policy. 2013;31:149-56.
75. Overdyk LM, Braid HE, Naaum AM, Crawford SS, Hanner RH. Real-time PCR identification of lake whitefish *Coregonus chupeaformis* in the Laurentian Great Lakes. J Fish Biol. 2016;88(4):1460-74.
76. Cammà C, Di Domenico M, Monaco F. Development and validation of fast Real-Time PCR assays for species identification in raw and cooked meat mixtures. Food Control. 2012;23(2):400-4.
77. Paine MA, McDowell JR, Graves JE. Specific identification using COI sequence analysis of scombrid larvae collected off the Kona coast of Hawaii Island. Ichthyol Res. 2007;55(1):7-16.
78. Primer3Plus 2016 [Available from: <http://primer3plus.com/cgi-bin/dev/primer3plus.cgi>].
79. OligoAnalyzer 3.1 [Internet]. Integrated DNA Technologies. 2016. Available from: <https://www.idtdna.com/calc/analyzer>.
80. Scott WB, Crossman EJ. Freshwater fishes of Canada. Fisheries research Board of Canada Bulletin. 1973;184.
81. Todd TN, Stedman RM. Hybridization of Ciscoes (*Coregonus* spp.) in Lake Huron Can J Zool. 1989;67:1679-85.
82. Goto A, Yokoyama R, Sideleva VG. Evolutionary diversification in freshwater sculpins (Cottoidea): a review of two major adaptive radiations. Environ Biol Fishes. 2014;98(1):307-35.

83. Froese R. *Myoxocephalus thompsonii* (Girard, 1851) Deepwater Sculpin 2016 [Available from: <http://www.fishbase.ca/summary/4124>.
84. Froese R. *Myoxocephalus quadricornis* (Linnaeus, 1758) Foruhorn sculpin 2016 [Available from: <http://www.fishbase.org/Summary/speciesSummary.php?ID=4122&AT=Hornsimpa>.
85. Eschmeyer PH, Bailey RM. The Pygmy Whitefish, *Coregonus Coulteri*, in Lake Superior. Trans Am Fish Soc. 1955;84(1):161-99.
86. Rodriguezlazarro D. A rapid and direct real time PCR-based method for identification of *Salmonella* spp. J Microbiol Methods. 2003;54(3):381-90.
87. Mieog JC, van Oppen MJH, Cantin NE, Stam WT, Olsen JL. Real-time PCR reveals a high incidence of Symbiodinium clade D at low levels in four scleractinian corals across the Great Barrier Reef: implications for symbiont shuffling. Coral Reefs. 2007;26(3):449-57.
88. Koek AG, Bruisten SM, Dierdorp M, van Dam AP, Templeton K. Specific and sensitive diagnosis of syphilis using a real-time PCR for *Treponema pallidum*. Clin Microbiol Infect. 2006;12(12):1233-6.
89. Bio-Rad. Real-Time PCR Applicaitons Guide. B ed: Bio-Rad Laboratories; 2006. p. 1-105.
90. Bowers HA, Tengs T, Jr. HGG, Burkholder JM, Rublee PA, Oldach DW. Development of Real-Time PCR Assays for Rapid Detection of *Psiesteria piscicida* and Related Dinoflagellates. Appl Environ Microbiol. 2000;66(11):4641-8.
91. Reed JZ, Tollit DJ, Thompson PM, Amos W. Molecular scatology: the use of molecular genetic analysis to assign species, sex and individual identity to seal faeces. Mol Ecol. 1997;6:225-34.
92. Chakraborty C, Doss CG, Patra BC, Bandyopadhyay S. DNA barcoding to map the microbial communities: current advances and future directions. Appl Microbiol Biotechnol. 2014;98(8):3425-36.
93. Krishna Krishnamurthy P, Francis RA. A critical review on the utility of DNA barcoding in biodiversity conservation. Biodivers Conserv. 2012;21(8):1901-19.
94. Zhang J, Hanner R. DNA barcoding is a useful tool for the identification of marine fishes from Japan. Biochem Syst Ecol. 2011;39(1):31-42.
95. Shendure J, Ji H. Next-generation DNA sequencing. Nat Biotechnol. 2008;26(10):1135-45.
96. Parson W, Pegoraro K, Niederstatter H, Foger M, Steinlechner M. Species identification by means of the cytochrome b gene. Int J Legal Med. 2000;114:23-8.
97. Hseih H-M, Chiang H-L, Tsai L-C, Lai S-Y, Huang N-E, Linacre A, et al. Cytochrome b gene for species identification of the conservation animals. Forsenic Science International. 2001;122:7-18.
98. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010;11(1):31-46.